Custom-designed methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) test for profiling of promoter methylation in ovarian and endometrial cancers

University of Helsinki

Faculty of biological and Environmental Sciences

Department of Bioscience

Division of Genetics

Master's thesis

Soheila Akhondzadeh

August 2016

"Raise your words, not voice. It is rain that grows flowers, not thunder."

Jalāl ad-Dīn Muhammad Rūmī

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta – Fakultet – Faculty	Laitos – Institution– Department
Faculty of Biological and Environmental Sciences	Department of Biosciences

Tekijä – Författare – Author: Soheila Akhondzadeh

Työn nimi – Arbetets titel – Title

Custom-designed methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) test for profiling of promoter methylation in ovarian and endometrial cancers

Oppiaine – Läroämne – Subject: Genetics

Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä – Sidoantal – Number of pages
Master's Thesis	August 2016	67 pages

Tiivistemä – Referat – Abstract

Background: Epithelial ovarian cancer is the most common type of ovarian cancer and is the most lethal gynecologic cancer due to its late diagnosis. Compared to ovarian cancer, endometrial carcinoma, as the most common gynecologic malignancy, is referred to as the "curable cancer", as it can be detected early. As aberrant promoter methylation patterns are a common change in human cancer, detection of promoter methylation status may help in early diagnosis. In this study, we used a custom-designed methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) assay as a rapid and easy method, to simultaneously detect the methylation status of multiple genes in ovarian and endometrial cancer samples.

Aims: To design and test an MS-MLPA assay for analyzing promoter methylation of four genes associated with ovarian and endometrial cancers. The selected genes were HNF1 homeobox B (HNF1β), Ten-eleven translocation 1(TET1), L1 cell adhesion molecule (L1CAM), and AT-rich interactive domain 1A (ARID1A). These genes are known to have expression changes by DNA methylation.

Methods: The promoter DNA methylation patterns of these four genes were analyzed in 15 cancer cell lines and 5 normal cell lines and DNAs using bisulfite sequencing. Six synthetic probe pairs were designed and optimized by applying them to cancer and normal cell lines and normal DNAs and comparing the results with those of bisulfite sequencing. Finally, the MS-MLPA assay was performed on patient specimens according to the MRC-HOLLAND MS-MLPA general protocol and methylation frequencies were calculated from MS-MLPA data.

Results and conclusion: The MS-MLPA assay gave accurate methylation results with the 170 samples assayed. The HNF1B, L1CAM, and TET1 Genes were observed methylated in tumor samples whereas they were not methylated in the normal samples or showed very little methylation, suggested to be favorable diagnostic markers. MS-MLPA robustly and sensitively detects the promoter DNA methylation status.

Avainsanat – Nyckelord – Keywords

Ovarian cancer, Endometrial cancer, DNA methylation, MS-MLPA, Bisulfite sequencing

Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors

Professor Päivi Peltomäki and M.Sc. Anni Niskakoski

Säilytyspaikka – Förvaringställe – Where deposited: Department of Bioscience, Division of Genetics

Muita tietoja – Övriga uppgifter – Additional information

Performed at Biomedicum, Department of Medical and clinical Genetics, University of Helsinki

CONTENTS

ABBREVIATIONS

Alu	Recombinogenic sequence
ARID1A	AT-rich interactive domain 1A
CIMP	CpG island methylator phenotype
CIN	Chromosomal instability
COBRA	Combined bisulfite restriction analysis
CpG	Cytosine-phosphate-guanine dinucleotide
CRC	Colorectal cancer
Dm	Methylation dosage ratio
DNMT	DNA (cytosine-5) methyltransferase
EC	Endometrial cancer/carcinoma
FFPE	Formalin-fixed paraffin-embedded
HNF1B	Hepatocyte nuclear factor-1beta
HNPCC	Hereditary non-polyposis colorectal cancer
IHC	Immunohistochemistry
L1CAM	L1 cell adhesion molecule
LHS	Left hybridization sequence
LINE-1	Long interspersed nuclear element 1
LOH	Loss of heterozygosity
LPO	Left probe oligonucleotide
LS	Lynch syndrome
MeDIP	Methylated DNA immunoprecipitation
MLH1	Human mutL homolog 1 (E. coli)
MLH3	Human mutL homolog 3 (E. coli)
MSH2	Human mutS homolog 2 (E. coli)
MSH6	Human mutS homolog 6 (E. coli)
MSI	Microsatellite instability
MS-MLPA	Methylation-specific multiplex ligation-dependent probe amplification
MSP	Methylation-specific PCR
RHS	Right hybridization sequence
RPO	Right probe oligonucleotide
Sporadic	Appearing in isolated instances
Tm	Annealing temperature
TET1	Ten-eleven translocation 1
TKF	Human female DNA from multiple donors
TSG	Tumor suppressor gene

Introduction

Epithelial ovarian cancer is the most lethal cancer among gynecologic cancers. The reason for this is the lack of symptoms in the early stage. Epithelial ovarian cancers are morphologically heterogeneous. Based on the cell type, ovarian cancers are divided into serous, mucinous, endometrioid, and clear cell (Prat, 2012). Compared to ovarian cancer, endometrial carcinoma (EC), as the most common gynecologic cancer, is highly curable due to its early diagnosis. Like ovarian cancer, the majority of endometrial cancers are sporadic and only 5% to 10% are familial. In endometrial carcinoma, Lynch syndrome is the most common hereditary cause, while in ovarian cancer, Lynch syndrome is the third common hereditary reason after BRCA1 and BRCA2 mutations (Ryan et al., 2005; Lynch et al., 2009).

Since alteration in the DNA methylation status is known to be common in human cancer, DNA methylation pattern can be used as a diagnostic marker for cancer. One purpose of the present study was to detect the promoter DNA methylation patterns of four genes in clinical samples of sporadic ovarian cancer and precursor lesions of endometrial carcinoma. We aimed to study whether these genes can act as diagnostic markers for ovarian and endometrial cancers.

Alterations in DNA methylation can be in the form of hypermethylation or hypomethylation. Hypermethylation can silence tumor suppressor genes and tumor suppressive microRNAs, while the activation of oncogenes and oncogenic miRNAs can be the result of hypomethylation (Sharma et al., 2010). For DNA methylation analysis, we used the methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) assay. Unlike MS-MLPA, most common methods for DNA methylation analysis (MSP, qMSP, COBRA and bisulfite sequencing) are bisulfite conversion dependent, time and work consuming, and not suitable for paraffin-embedded tissue samples. The MS-MLPA test, as a fast and easy method, can be applied successfully to paraffin-embedded tissue samples, which usually have poor quality (Nygren et al., 2005).

Since a commercial MS-MLPA kit is available for a few genes known to be methylated in ovarian and endometrial carcinoma, a custom-made MS-MLPA test was designed to

analyze the methylation patterns of four genes associated with these carcinomas. Another purpose of this study was to validate the accuracy of the custom-designed MS-MLPA assay when applying to paraffin-embedded tissues. For testing optimum performance of custom-made test in DNA methylation detection, we applied the MS-MLPA test to the cancer and normal cell lines and healthy control DNA samples and compared the findings with bisulfite sequencing results.

The selected genes were HNF1 homeobox B (HNF1β), Ten-eleven translocation 1(TET1), L1 cell adhesion molecule (L1CAM), and AT-rich interactive domain 1A (ARID1A). In L1CAM and ARID1A genes, two islands were studied. First, the promoter DNA methylation status of these genes were analyzed in 20 cancer and normal cell lines and DNAs using bisulfite sequencing and the MS-MLPA assay. As both results were comparable, the custom-designed MS-MLPA test was performed on ovarian and endometrial cancer samples.

Review of the literature

1. Cancer

Cancer has traditionally been considered a group of related diseases that develop due to the accumulation of genetic mutations (Hanahan &Weinberg, 2000). However, the importance of cancer epigenetics has been increasingly highlighted recently. Six hallmarks are considered for cancer by Hanahan and Weinberg (2000). Maintaining chronic proliferation is the first hallmark of cancer. Normal tissues undergo controlled proliferation, while abnormal proliferation in cancer cells results in the development of tumors. As the second hallmark of cancer, due to the inactivation of tumor suppressor genes, cells continue to grow and proliferate without any limitations which can result in evolving cancer. The third hallmark of cancer is evasion of apoptosis. Programmed cell death is required to prevent cancer development. Cancer cells with resistance to programmed cell death can succeed in progressing to tumors. As another hallmark of cancer, unlimited replication is required for cancer cells to produce tumors. An ability to induce blood-vessel formation is required for cancer cells, not just to supply oxygen and nutrients, but to remove metabolic waste products and carbon dioxide. Tissue invasion and metastasis, the sixth hallmark of cancer, is the spread of cancer to another organ (Hanahan & Weinberg, 2000).

In 2011, Hanahan and Weinberg added two new hallmarks to their list, including an abnormal energy pathway and evading immune surveillance. Normal tissues, under normal conditions, utilize aerobic oxidative phosphorylation to generate energy, while cancer cells undergo aerobic glycolysis to produce energy. This glycolytic switch can result in the activation of oncogenes and mutation in tumor suppressor genes. Evading immune surveillance is the eighth hallmark of cancer. The immune system plays an important role in resisting the formation of tumors. Cancer cells, by escaping immune surveillance can succeed in developing to tumors. Genomic instability and chronic inflammation are the underlying hallmarks that enable these eight hallmarks to develop and result in tumor progression (Hanahan & Weinberg, 2011).

1.1. Cancer genetics

Cancer can arise from the accumulation of genetic alterations including base substitutions, deletions, insertions, amplifications, alterations in chromosome number, and chromosome translocations (Lengauer et al., 1998). These alterations are mostly sporadic. Three types of cancer related genes, including oncogenes, tumor suppressor genes, and stability genes, are responsible for tumorigenesis (Vogelstein & Kinzler, 2004).

1.1.1. Oncogenes, tumor suppressor genes, and DNA mismatch repair genes

Proto-oncogenes play a key role in regulating cell growth and differentiation. When proto-oncogenes are mutated or highly expressed, they can predispose to cancer genes called oncogenes. Oncogene activation by stimulation of cell birth or the inhibition of cell death can lead to tumorigenesis. One somatic mutation is sufficient for the activation of oncogenes. Oncogenes can be activated due to chromosomal translocation, gene amplification, and subtle intragenic mutation. Tumor suppressor genes are responsible for inhibiting cell growth. Unlike oncogenes, if they are not expressed because of a mutation, increased proliferation and tumor development result. According to the two-hit hypothesis, inactivation of both alleles in a tumor suppressor gene is essential for tumorigenesis (Vogelstein & Kinzler, 2004).

DNA mismatch repair (MMR) genes are responsible for recognizing and repairing the base-base mismatches and insertion/deletion mispairs that occur during DNA replication. Of the MutS-related proteins, MSH2-MSH3 and MSH2-MSH6 complexes are required for mismatch recognition, while MLH1-PMS2 and MLH1-MLH3 are two complexes of MutL-related proteins responsible for recognizing mismatches. Germline mutations in mutL homologue 1 (MLH1), mutS homologue 2 (MSH2), MSH6, or postmeiotic segregation increased 2 (PMS2) can cause Lynch syndrome. Like tumor suppressor genes, mutations in two alleles are required for defective mismatch repair (Peltomäki, 2014).

1.1.2. Genomic instability in cancer

Genomic instability as a characteristic of cancer has various forms, including chromosome instability (CIN) and microsatellite instability. In chromosome instability, chromosome structures and numbers change over time in a cancer cell population (Negrini et al., 2010). Microsatellite instability (MSI) is caused by deficiencies in DNA mismatch repair (MMR) genes. The mismatch repair pathway is responsible for correcting replication errors that occur during DNA replication in microsatellite sequences (Sameer et al., 2014).

1.1.2.1. Chromosome instability (CIN)

Since the majority of cancers display chromosome instability (CIN), it has been considered a predominant hallmark of cancer. The definition of chromosome instability is the elevated rate of changed in karyotypes in a cell population. CIN, as the main form of genomic instability, can be divided into numerical CIN and structural CIN. Numerical CIN is the gain or loss of a whole chromosome, causing aneuploidy, while Structural CIN is abnormality in structure of chromosomes, resulting in translocation, deletions, and amplifications of DNA. In practice, due to several reasons, measuring CIN is difficult. The loss of accurate molecular methods to profile cell to cell variability, difficulty in detecting chromosomal aberrations over time, and the diversity of chromosomal aberrations are the challenges in measuring CIN (Heng et al., 2013).

CIN plays an important role in cancer because it can either contribute to cancer or suppress tumor development. In hereditary cancers, CIN can cause cancer development by inducing mutations in DNA repair genes. In sporadic cancer, two models can explain the role of CIN: the mutator hypothesis and oncogene-induced DNA replication stress model. The mutator hypothesis emphasizes the rarity of mutation in normal cells and the frequency of spontaneous mutations in cancer cells. Although in sporadic cancer the molecular basis of genomic instability is not clear, the latter model states that oncogene activation can cause deregulated DNA replication (Negrini et al., 2010).

1.1.2.2. Microsatellite instability (MSI*)*

Microsatellite instability (MSI) is an alteration of the length in repetitive genomic sequences due to insertions or deletions and is caused by mutations in DNA mismatch repair (MMR) genes (Sameer et al., 2014). MMR is as one of the repair mechanisms that recognizes and repairs small DNA loops. Mutated MMR genes can fail in repairing errors in repetitive sequences, which results in MSI. Although MSI can be seen in tumors of different organs, it is considered as a marker of Lynch syndrome and colorectal cancer (Kurzawski et al., 2004). MSI caused by deficient MMR genes is a hallmark of LS tumors but it is also detected in about 15% of sporadic CRCs (Dionigi et al., 2007).

1.2. Cancer epigenetics

Epigenetics has been defined as ‘‘the study of changes in gene function that are mitotically and/or meiotically heritable and that do not entail a change in DNA sequence’’ (Riggs & Porter, 1996). Epigenetic abnormalities play an important role in the promotion of cancer development, in addition to genetic alterations (Jones & Baylin, 2007). Epigenetic mechanisms can be divided into four categories: DNA methylation, histone modification, noncoding RNAs, including microRNAs (miRNAs), and chromatin remodeling (Wong et al., 2007, Wilson & Roberts, 2011). The interplay of these mechanisms regulates the genome by regulating its accessibility and compactness (Sharma et al., 2010).

1.2.1. DNA methylation

DNA methylation as the best known epigenetic mechanisms plays a key role in the regulation of gene transcription. Alterations in DNA methylation in the form of site-specific hypermethylation or genome-wide hypomethylation can contribute to carcinogenesis (Das and Singal, 2004). Although about 70% of all CpG dinucleotides are methylated in the human genome, CpG dinucleotides within CpG islands are mostly unmethylated during development and in normal tissues (Kim et al., 2009). Methylation of CpG islands has been identified as the gene silencing mechanism that has a key role in regulation of gene expression and chromatin structure. In the mammalian genome, DNA methyltransferases (DNMTs) is responsible for adding a methyl group to the position 5 carbon on the cytosine (C) nucleotide, most commonly at the 5’-CG-3’ dinucleotides. Several DNMTs have been identified, including: DNMT1, DNMT1b, DNMT1o, DNMT1p, DNMT2, DNMT3A, DNMT3b with its isoforms, and DNMT3L. The other mechanisms involved in DNA methylation are demethylases, methylation centers, and methylation protection centers (Das and Singal, 2004).

Hypomethylation, which is a loss of methylation in the 5-methylcytosine nucleotide, has typically been found in all neoplasms. Hypomethylation can contribute to carcinogenesis by several mechanisms. First, hypomethylation in repetitive sequences, such as satellite

or pericentromic regions, can result in genomic instability. In addition, hypomethylation of retrotransposons can result in abnormal gene structure and function. Hypomethylation can also contribute to the activation of oncogenes. Finally, hypomethylation can lead to a loss of imprinting, resulting in gene expression imbalances and cancer development. Hypermethylation, as an important event in developing cancer, can lead to the repression of transcription in tumor suppressor genes, mismatch repair genes, and cell-cycle-regulatory genes at CpG islands and contribute to gene silencing. It also can result in a loss of imprinting in cancer (Wong et al., 2007).

1.2.2. Histone modification, Chromatin remodeling and microRNA

Chemical modifications of histone such as methylation, acetylation, phosphorylation, and ubiquitinylation, can modify the chromatin structure from inactive heterochromatin to active euchromatin, and vice versa (Wong et al., 2007). Depending upon the type of residues and chemical modifications, histone modifications can either activate or repress gene promoters. For example, histone methylation can either have repressive or activating effect on chromatin depending on the methylated amino acid residue, while histone acetylation usually activate transcription (Ferrari et al., 2014). Nucleosome remodeling, as a non-covalent mechanism, has a key role in the regulation of gene activity. Mutations in remodelers, which use ATP hydrolysis to alter chromatin structure, can cause cancer development. MicroRNAs (MiRNAs) are small non-coding endogenous RNAs that not only regulate gene expression at a post-transcriptional level, but also play a key role in epigenetic mechanisms. Epigenetic modifications can affect microRNAs expression, and microRNAs also have the capacity to control the epigenome (Sharma et al., 2010).

2. Lynch syndrome

Lynch syndrome, which is traditionally known as hereditary nonpolyposis colorectal cancer (HNPCC), is an autosomal dominant disorder caused by inherited aberration in the DNA mismatch repair (MMR) system and is associated with tumors with microsatellite instability (MSI). People with Lynch syndrome have a high risk of developing colorectal, endometrial, ovarian and some other cancers (Aarnio et al., 1999). Mismatch repair genes, including MSH2, MSH6, MSH3, MSH4, MSH5, MLH1, PMS2, PMS1, and MLH3, are responsible for repairing mistakes made during DNA replication in humans. Inherited defects in MMR genes, mutL homologue 1 (MLH1), mutS homologue 2 (MSH2), MSH6, or postmeiotic segregation increased 2 (PMS2) can cause Lynch syndrome (Lynch et al., 2009).

DNA MMR genes can be inactivated by either genetic or epigenetic mechanisms. Inactivation of both alleles in DNA MMR genes is required for tumor development. In hereditary cancer, mutation in one allele is inherited and the second allele is inactivated in the somatic cell, while in sporadic cancer, inactivation of both alleles occurs somatically. Epigenetic mechanisms play a fundamental role in Lynch syndrome by inactivating tumor suppressor genes or causing epimutations in MMR genes. MMR genes can be inactivated by epigenetic events, including constitutional epimutations as the first hit and promoter methylation as the second hit (Peltomäki, 2014).

The first diagnostic criteria for Lynch syndrome diagnosis was the Amsterdam criteria. According to the classic Amsterdam criteria, a Lynch syndrome family contains three or more individuals with a Lynch syndrome-associated cancer in 2 generations at an age of less than 50 years (Stoffel et al., 2009). The Bethesda guidelines were developed in 1997 and revised in 2003 for the genetic diagnosis of Lynch syndrome. The Bethesda guidelines for Lynch syndrome recommend microsatellite instability (MSI) testing or abnormal immunohistochemistry (IHC) of the tumors from individuals with HNPCC. According to several studies, the Bethesda guidelines are the most accurate (94%) guidelines, followed by the Amsterdam criteria II (72%) and classic Amsterdam criteria (61%) (Umar et al., 2004).

As many Lynch syndrome families do not fulfill Amsterdam criteria or Bethesda guidelines, or they meet these two guidelines but they do not have germline mutations in any DNA MMR genes, the *Jerusalem criteria* were developed in 2010. According to these guidelines, all CRCs diagnosed at an age of 70 or younger, should be screened by MSI or IHC testing (Boland & Lynch., 2013).

Since the basic epigenetic patterns are specific to tissue types, methylation changes observed in Lynch tumors are essentially expected to resemble those seen in the corresponding sporadic tumors (Nieminen et al., 2016). Therefore, knowledge of epigenetic patterns obtained by studying Lynch-associated endometrial and ovarian tumors is likely to be valid for their sporadic counterparts and vice versa. This aspect will be further discussed in later sections (e.g. Conclusions and future prospects, p.59)

3. Ovarian cancer

Ovarian cancer is the sixth most common cancer in women. It accounts for 6.1 cases per 100, 000 women and its lifetime risk is 0.5 %(Bogliolo et al., 2016). Ovarian cancer is the most lethal gynecological cancer, as it causes over 100, 000 deaths per a year among women worldwide (Prat, 2012). The reason for the high mortality rates is the late diagnosis in most cases (Jayson et al., 2014). Most cases of ovarian cancer are sporadic and only 5% to 10% of cases are familial. There are different risk factors for ovarian cancer. A family history of the disease, BRCA1 and BRCA2 mutations, and Lynch syndrome are the most significant risk factors for ovarian cancer (Holschneider & Berek, 2000).

Epithelial ovarian cancer is a heterogeneous cancer that is classified into five main groups based on histopathology and molecular genetics. These five distinct types are high-grade serous, endometrioid, clear cell, mucinous, and low-grade serous. High-grade and low-grade serous are different tumor types, and are the most and the least common types of ovarian cancer, respectively (Prat, 2012). In another classification, ovarian cancer is classified into type I and II tumors. Type I tumors, which are low grade, include endometrioid, mucinous, and clear cell type. Type I tumors contain mutations in BRAF, KRAS, and PTEN with microsatellite instability. Type II tumors are high grade, including high-grade serous and carcinosarcoma. Mutations in p53, BRCA1, and BRCA2 are frequently detected in type II tumors (Jayson et al., 2014).

High-grade serous carcinomas (HGSCs), as the most common and lethal types of ovarian cancer, are characterized by p53 gene abnormalities and deleterious mutations in BRCA1 and BRCA2. Additionally, genomic instability and DNA copy number abnormalities have been detected in the most high-grade serous cancer. Low grade serous carcinomas (LGSC) are uncommon and exhibit KRAS and BRAF mutations. Unlike HGSCs, LGSCs are not related with BRCA1 and BRCA2 germline mutations. It has been proposed that both HGSC and LGSC arise from distal fallopian tube (Prat, 2012).

Mucinous ovarian cancers have KRAS mutations and a high frequency of HER2 amplification (Jayson et al., 2014). In clear-cell carcinomas (CCCs), mutations in the ARID1A gene are prevalent (Prat, 2012). Upregulation of HNF1B has been identified in CCCs (Kato et al., 2007). Endometrioid ovarian cancers account for 10% of all ovarian carcinomas. This carcinoma has been reported more in patients with Lynch syndrome compared to other types of ovarian cancer. ARID1A mutations and somatic mutations of the beta-catenin (CTNNB1) and PTEN genes are the most prevalent abnormalities in endometrioid carcinomas. Most CCCs and ECs seem to originate from ovarian endometriosis (Prat, 2012).

3.1. Lynch syndrome-associated ovarian cancer

In women with Lynch syndrome, ovarian cancer can be the presenting cancer or a second cancer. Women with Lynch syndrome have a higher risk (6-8%) of developing ovarian cancer than the 1.4% risk in the general population. Compared to the general population, a younger diagnosis age (42-48) has been reported for Lynch syndrome-associated ovarian cancer. Lynch syndrome accounts for 10-15% of hereditary ovarian cancer cases. Unlike sporadic ovarian cancer, which is mostly diagnosed in stage III or IV, 80% of Lynch syndrome-associated ovarian cancers are diagnosed in stages I or II. Histologically, ovarian cancers in Lynch syndrome have a non-serous type (Nakamura et al., 2014). Ketabi et al. (2011) focused 0n 63 Swedish and Danish families with Lynch syndrome-associated ovarian cancer and found a variety of histopathological subtypes including, endometrioid (35%), serous (28%) and clear cell (17%). Bonadona et al. (2011) studied 537 French families with Lynch syndrome and found a 20% lifetime risk of ovarian cancer for MLH1 mutation carriers, 24% for MSH2 and 1% for MSH6. They concluded that the lifetime risks for ovarian cancer is mostly associated with MLH1 or MSH2 mutations.

4. Endometrial cancer

Endometrial carcinoma (EC) is the most common gynecologic cancer and it accounts for 15-20 cases per 100, 000 women per year. The lifetime risk for this cancer is 2.7% and the median age of diagnosis is 65 years. Endometrial cancer is highly curable due to its early diagnosis. Most of the endometrial cancers (90%) are sporadic and only 10% are familial (Ryan et al., 2005). The majority of familial endometrial cancer are associated with hereditary non-polyposis colorectal cancer (HNPCC). The risk factors for endometrial cancer are family history, estrogen exposure, older age, obesity, early menarche, late menopause and infertility (Burke et al., 2014).

Sporadic endometrial cancers are morphologically heterogeneous and are divided into two groups based on histological features; estrogen-dependent adenocarcinoma (type I) with endometrioid morphology and estrogen-independent carcinoma (type II) with serous papillary and clear cell morphology. The first group, which are more common (90% of ECs), are tumors that occur in pre- and post-menopausal women and have a good prognosis. Type I tumors are usually low grade and are associated with abnormalities in DNA-mismatch repair (MMR) genes, K-ras, PTEN and beta-catenin. The second group consists of tumors that occur in post-menopausal women and they usually show highly aggressive clinical behaviors. Type II tumors are commonly associated with abnormalities of p53 and HER2/neu (Ryan et al., 2005).

For both of type I and type II endometrial adenocarcinoma, precursor lesions are known: endometrial hyperplasia, which is the precursor of endometrioid adenocarcinoma, and endometrial intraepithelial carcinoma (EIC), which is the precursor of serous and clear cell carcinomas. Endometrial hyperplasia can be classified into simple hyperplasia with and without cytologic atypia, and complex hyperplasia with and without cytologic atypia. Endometrial hyperplasia with atypia is the least common type but is the type most likely to develop to endometrial (Type I) carcinoma. Simplex hyperplasia rarely develop to endometrial carcinoma (Wang et al., 2015).

4.1. Lynch syndrome-associated endometrial cancer

The majority of inherited endometrial carcinomas are associated with Lynch syndrome. Women with Lynch syndrome with MLH1 or MSH2 mutation has a risk of 40% of developing endometrial cancer. The median age at which women with Lynch syndrome are diagnosed with endometrial cancer is 49 years. Lynch syndrome-associated endometrial cancers primarily have an endometrioid histology, but other types, including clear cell, papillary serous and MMMT, have been reported (Nakamura et al., 2014). Bonadona et al. (2011), by studying 537 French families with Lynch syndrome, found a 54% lifetime risk of endometrial cancer for MLH1 mutation carriers, 21% for MSH2 and 16% for MSH6. In another study of 155 women with Lynch syndrome-associated endometrial cancer, Stoffel et al. (2009) found a lifetime risk of 39.39% for MLH1 or MSH2 mutation carriers. These results indicated that lifetime risks for endometrial cancer are mostly associated with MLH1 or MSH2 mutations.

5. Methods to study methylation

There are different methods for studying DNA methylation as a common epigenetic change in mammals. These approaches can be divided into two main groups: global DNA methylation analysis and gene-specific methylation analysis. High-performance liquid chromatography (HPLC) is a highly quantitative method to quantify global DNA methylation. In this method DNA is first hydrolyzed and, then, the 5 methylated cytosines (5 mC) and deoxycytidine (dC) are separated chromatographically and the methylation levels are measured by comparing the relative absorbance of cytosine and methylasytosine. This method is limited by requiring a large amount of DNA and not being suitable for high-throughput analysis (Fraga & Esteller, 2002). Global DNA methylation can also be assessed by bisulfite-based PCR methods. These methods are based on bisulfite treatment of DNA and PCR products of repetitive DNA elements, such as Alu elements and long interspersed nucleotide elements (LINE). Unlike HPLC, bisulfite-based PCR methods require very little DNA. Methylated DNA immunoprecipitation (MeDIP) is a genome-wide method for the determination of methylation patterns. In this high-resolution method, anti-methylcytosine is used to immunoprecipitate DNA containing highly methylated CpG sites. Then, immunoprecipitated DNA is hybridized to a microarray. Like HPLC, a large amount of genomic DNA is required in this method (Shen& Waterland, 2007).

There are several gene-specific approaches to studying DNA methylation. In methylation-specific PCR (MSP), PCR is performed by using two sets of primers that are able to amplify either methylated or unmethylated alleles. MSP is a highly sensitive method that can even be applied to paraffin-embedded tissues with poor quality, but is not a quantitative method. Quantitative methylation-specific PCR (qMSP) is a quantitative method of MSP, which is frequently used to detect the amount of methylation at specific loci. Compared with MSP, in this method, a reference sequence in a separate reaction is required for assay optimization (Shames et al., 2007). Combined bisulfite restriction analysis (COBRA) is a quantitative technique for methylation detection. In this method, after bisulfite treatment and PCR amplification, PCR products are treated with a restriction enzyme and digested fragments are finally separated by polyacrylamide gel

electrophoresis. Unlike MSP, this method is time consuming and the analysis can only be done for specific restriction enzyme cutting sites. Bisulfite-pyrosequencing is another method for DNA sequencing. In Bisulfite-pyrosequencing chemical light reaction is used to detect the incorporation of nucleotides into the single stranded DNA. A molecule of pyrophosphate (PPi) is released from incorporated nucleotides into the strand, and converted to ATP. ATP provides the energy for a light reaction in which luciferin is converted to oxyluciferin by luciferase. Light is captured and visualized graphically (Shames et al., 2007). Illumina Infinium HumanMethylation 450 BeadChip (450K) is an array based profiling approach for DNA methylation detection on a genome-wide scale. This array contains480, 000 probes, and targets 99% of genes and 96% of CpG islands (Marabita et al., 2013). Bisulfite sequencing developed by Frommer et al (1992) is regarded as a gold-standard technology for profiling the DNA methylation status. In this methods, after bisulfite modification, unmethylated cytosines are converted to uracil whereas methylated cytosines remain unaffected. In subsequent PCR, the region of interest is amplified in both methylated and unmethylated alleles. After PCR amplification, DNA can be sequenced to quantify DNA methylation status (Li et al., 2011). This method is described in more detail below (Figure 1 and page 36). As most methods for analyzing DNA methylation status are based on bisulfite conversion and require a great amount of high-quality DNA, Methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) was developed to analyze methylation status and copy number changes of genomic DNA (Nygren et al., 2005).

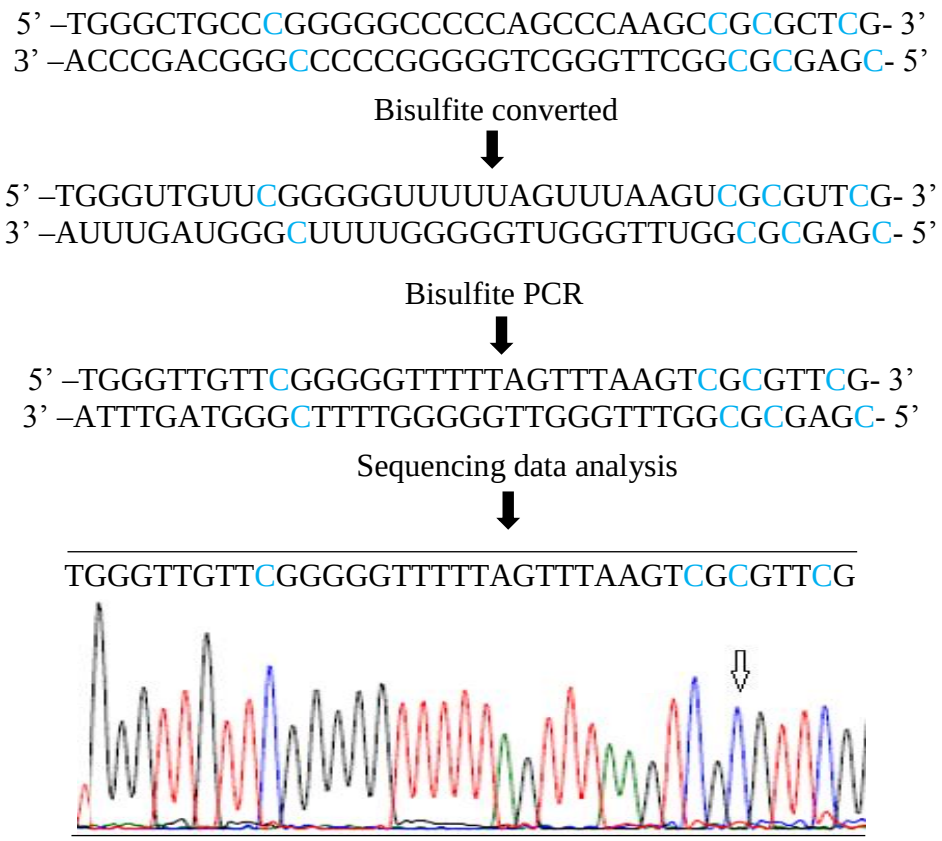


Figure1. Outline of bisulfite sequencing. In bisulfite conversion of genomic DNA, unmethylated cytosine residues (black) are converted to uracil, while methylated cytosines (blue) remain unaffected. In subsequent PCR, uracil residues are converted to thymine. DNA methylation pattern can be detected by PCR sequencing.

5.1. The MS-MLPA method

Methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) is a method to detect simultaneously changes in the gene copy number and methylation status of up to 40 chromosomal sequences using only 20 ng of genomic DNA. MS-MLPA probes contain a methylation sensitive HhaI recognition site (GCGC). In this method, after genomic DNA denaturation and MS-MLPA probes hybridization, the probe-DNA complex is ligated and digested by a methylation-sensitive enzyme. If the CpG site is methylated, a normal MS-MLPA signal will be produced. If the CpG island is not methylated, the genomic DNA-MS-MLPA probe complex will be digested by Hha1 and no amplification will take place. The major advantage of this method compared with other methods, such as MSP and bisulfite sequencing, is that it is not based on bisulfite conversion of unmethylated cytosines. Moreover, this very quantitative method allows for a combined copy number and DNA methylation detection in a single reaction. Furthermore, MS-MLPA can be performed on poor quality DNA, such as DNA from formalin fixed, paraffin-embedded tissues. Formaldehyde fixation and embedding in paraffin result in partial denaturation and degradation of the DNA, which do not influence the results in MS-MLPA test. Despite the many advantages, the MS-MLPA test has also been criticized. For example, methylation sites must contain a restriction site (GCGC) for the methylation-sensitive HhaI enzyme (Nygren et al., 2005). The principal of this method is described in more detail below in Figure 2 and page 39.

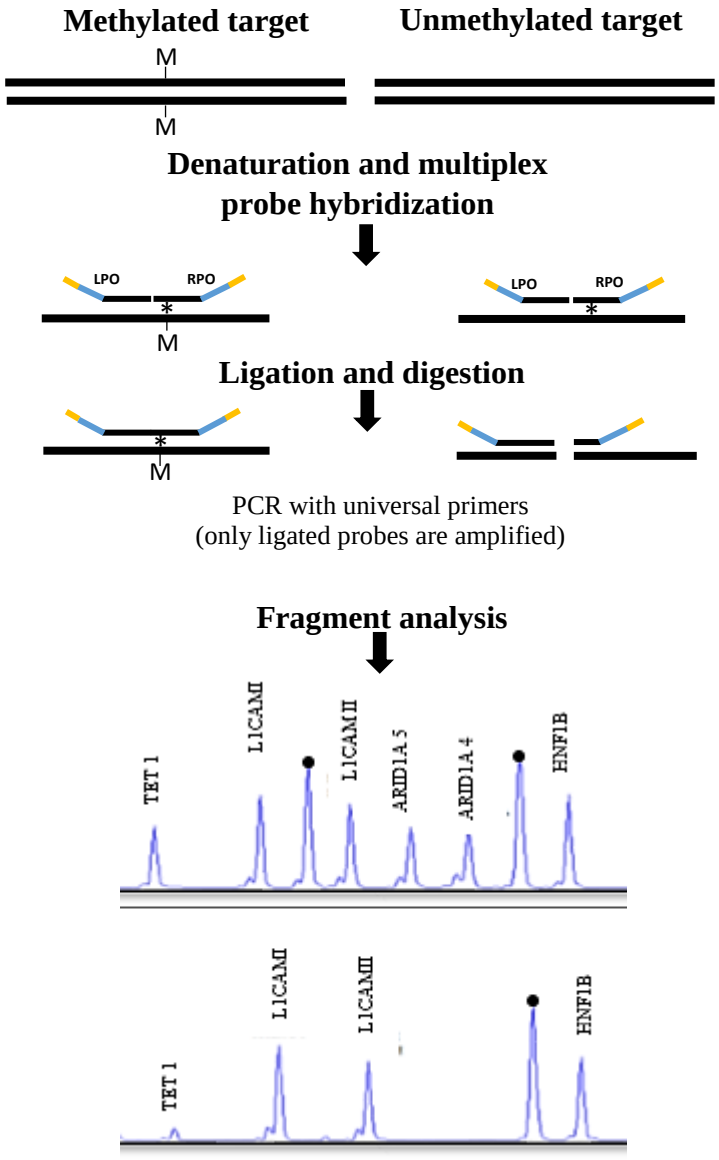


Figure 2. Outline of MS-MLPA procedure. MS-MLPA probes contain two oligonucleotides, Left probe oligonucleotide(LPO) and right probe oligonucleotide(RPO). Both oligonucleotides harbor universal primer sites (orange), optional stuffer sequence (blue), and hybridizing sequence (black). MS-MLPA probes contain a methylation-sensitive restriction site (GCGC= ∗) for HhaI. MS-MLPA Probe-DNA complex is simultaneously ligated and digested by methylation specific enzymes. If the CpG residues in the GCGC site is methylated, a normal MS-MLPA peak will be detected. If the CpG site is not methylated, MS-MLPA probe-DNA complex is digested by the HhaI and no signal will be detected. In the fragment analysis part, the top picture is the MS-MLPA results without HhaI treatment. The lower panel is the reults after digestion with HhaI. As it can be seen, there is no signal for ARID1A,4 and ARID1A,5 in the lower picture, indicating that the CpG site is not methylated in ARID1A. (● = control peaks).

6. Genes included in the custom MS-MLPA test

6.1. The HNF1B gene

Hepatocyte nuclear factor-1beta (HNF-1 beta), also known as transcription factor 2 (TCF2), is a member of the homeodomain-containing superfamily of transcription factors. This gene is located at chromosome 17q12 and composed of 9 coding exons. HNF1B has a major role in endodermal development (Nemejcova et al., 2016). In humans, HNF1B expression is tissue specific in normal tissues, which is most likely controlled at the epigenetic level (Kato et al., 2007). This gene has was observed to be an excellent marker for clear cell carcinoma (CCC) of the ovary (Tsuchiya et al., 2003).

Terasawa et al. (2006) studied HNF1B as a target for epigenetic inactivation in ovarian cancer. They examined the promoter methylation status of HNF1B in 15 ovarian cancer cell lines and 98 ovarian cancer specimens using combined bisulfite restriction analysis (COBRA). They found that eight (53%) of the ovarian cancer cell lines were heavily methylated at HNF1B promoter. In clinical ovarian cancer samples, HNF1B was methylated in 41.3% of serous samples, 25.0% of mucinous tumors, and 28.6% of endometrioid tumors. No methylation was detected in clear cell samples and normal ovarian tissues. They concluded that differences in the DNA methylation patterns between clear cell and other types of ovarian cancer may be a useful diagnostic marker.

To assess the correlation between the methylation and expression status, they used RT-PCR to check the gene expression of 15 ovarian cancer cell lines. Expression of HNF1B was not detected in any of the eight methylated cell lines, concluding a direct link between DNA methylation and gene silencing.

Cuff et al. (2013) characterized HNF1B promoter methylation in ovarian serous and clear cell histotypes. Among four ovarian serous cases, all of them exhibited methylated patterns, while in each of three ovarian clear cell carcinoma cases, no DNA methylation was detected. The relation between HNF1B promoter methylation and gene expression was previously reported in ovarian carcinoma (Terasawa et al., 2006). Cuff et al. (2013) concluded that as there is a direct link between DNA hypomethylation of HNF1B and its

expression in ovarian clear cell carcinoma, it can be used as a marker of clear cell phenotype.

Tsuchiya et al. (2003) analyzed 12,600 genes to identify genes associated with ovarian clear cell carcinoma, using the oligonucleotide array method. Sixteen genes were up regulated in CCC including HNF1B. They validated the over expression of HNF1B in mRNA and protein level using real-time quantitative reverse transcriptase-polymerase chain reaction and immunoblotting. Furthermore, immunohistochemical expression of HNF1B was analyzed in 83 ovarian cancer specimens including 22 CCC and 61 non-CCC specimens. Of 22 CCC cases, 21 showed immunostaining, while most non-CCC cases (60 of 61) were negative immunohistochemically. They concluded that the expression of HNF1B was strongly linked to ovarian clear cell carcinoma.

In another study, Shen et al. (2013) analyzed expression and DNA methylation patterns of HNF1B in serous ovarian cancer in The Cancer Genome Atlas (TCGA) Research Network. The results showed that HNF1B was epigenetically silenced in half the 576 serous ovarian tumors, while there was no evidence of methylation in normal fallopian tube samples.

To identify the relation between the methylation and expression status, they analyzed immunohistochemical expression of HNF1B in 1,149 ovarian cancers from the Ovarian Tumor Tissue Analysis Consortium, and DNA methylation patterns of 269 of these samples. Unmethylated, expressed HNF1B was detected in the majority of the CCC tumors, whereas the majority of serous tumors were immunohistochemically negative for HNF1B and showed frequent methylation of this gene. A direct link between the DNA methylation and protein expression was concluded.

Shen et al. (2013) also studied the promoter DNA methylation patterns of HNF1B in 32 serous and 4 clear cell ovarian tumors using the Illumina Infinium Human methylation 27 assay. The promoter region was methylated in 42% of serous cases and methylation of HNF1B was not detected in none of the clear cell ovarian tumors. This opposite manner of HNF1B in serous and CCC ovarian cases supported the hypothesis that HNF1B may

have a tumor suppressor role in serous cancer and an oncogenic role in clear cell ovarian cancer.

Nemejcova et al. 2016 analyzed HNF1B expression immunohistochemically in lesions of the endometrium in 320 cases including both tumor and non-tumor endometrial lesions. The expression of HNF1B was detected in 28% of endometrial endometrioid carcinoma (ECs), 26% of serous carcinoma, 83% of endometrial clear cell carcinoma, 95% of ovarian clear cell, 88% of hyperplasia with atypias, 91% of hyperplasias without atypias, and ≥ 80% of different normal endometrium samples.

They also examined the methylation patterns of HNF1B in 30 endometrial endometrioid cases, 19 ovarian clear cell carcinomas, and 15 corresponding normal endometrium tissues. Promoter DNA methylation was detected in 13.3% (4/30) of endometrioid ECs, whereas there was no evidence of DNA methylation in normal endometrium and in ovarian CCC. They concluded that the strong expression of HNF1B in CCCs compared to other subtypes can be helpful as a marker for ovarian CCC phenotype.

6.2. The L1CAM gene

The L1 cell adhesion molecule (L1CAM) is an axonal glycoprotein belonging to a large class of immunoglobulin superfamily cell adhesion molecules (CAMs) that is responsible for cell-to-cell adhesion. L1CAM is involved in the development of the nervous system (Kenwrick et al., 2000). The aberrant expression of L1CAM has been observed in many human tumors and is correlated with aggressive tumor behavior (Kato et al., 2009).

Kato et al. (2009) studied the L1CAM gene promoter methylation status and its protein expression pattern in 4 CRC cell lines and 71 primary CRCs. Of 4 CRC cell lines, 2 (SW480 and SW620) showed strong expression of L1CAM protein, whereas expression of L1CAM protein was not detected in two other CRC cell lines (Caco-2 and DLD-1). They also examined the expression of L1CAM in 71 patients with primary CRCs. Normal colonic epithelium were negative for the L1CAM expression and 31 (43.7%) of 71 carcinoma cases were histochemically positive. They concluded that L1CAM expression is regulated by epigenetic modification in CRC cell lines, and that DNA hypomethylation in the L1CAM CpG islands is correlated with L1CAM aberrant expression and contributes to aggressive behavior of tumors in CRC.

Huszar et al. (2009) analyzed the expression of the L1CAM in 272 endometrial endometrioid carcinoma cases and identified 29% of cases (78 of 272) were L1CAM positive and 71% (194 of 272) were L1CAM negative. They also examined the expression of the L1CAM in the mixed ECs cases composed of endometrioid and serous or clear-cell cases. The mixed ECs cases were L1CAM positive. These results suggested that normal endometrium and the majority of endometrioid ECs are L1CAM negative, whereas L1CAM was strongly expressed in serous and clear cell ECs. They concluded that immunohistochemical analysis of endometrioid ECs might be useful for the identification of patients with this type of tumors that have a bad prognosis.

Schirmer et al. (2013) examined the possible correlation between DNA methylation and protein expression of the L1CAM promoter in endometrial carcinomas (ECs) using RT-PCR, Western blotting and bisulfite sequencing. They investigated 6 endometrial carcinomas (ECs) cell lines and identified a low level of expression in 4 of 6 cell lines

(ECC1, HEC1A, EN1, MFE 296), while 2 of them were highly expressed (HEC1B and SPAC1L). They also studied the promoter methylation status of L1CAM in two promoter regions, one before the non-translated exon (exon 0) and one next to the first coding exon (exon 1). They identified that the cell lines with positive results in expression (HEC1B and SPAC1L) showed the lowest level of methylation in promoter 1, whereas a high level of methylation was detected in negative cell lines in expression. They concluded that promoter 1 methylation status is correlated with the level of L1CAM expression. They also studied the methylation of the L1CAM promoter in EC tumor tissues. They found a high variability in L1CAM promoter methylation and no marked differences were detected in promoter methylation of expression-positive and expression-negative samples.

In serous ovarian carcinoma, L1CAM expression in tumors and ascitic fluid associates with a highly invasive and metastatic propersity (Bondong et al., 2012). We are not aware of any published studies examining the relationship between L1CAM expression and methylation in endometrial hyperplasias or subtypes of ovarian carcinoma.

6.3. The TET1 gene

Ten-eleven translocation 1 (TET1), as a part of the TET protein family (TET1, TET2, and TET3), is a dioxygenase that plays an important role in cytosine demethylation. TET1 functions as a tumor suppressor gene in tumor cells (Hsu et al., 2012).

Ichimura et al. (2015) studied the DNA methylation and expression patterns of the TET1 gene in colorectal cancer, using bisulfite pyrosequencing, qRT-PCR, and western blot analysis. A high level of DNA methylation was detected in three of six colorectal cancer cell lines (SW480, RKO, and HT29), and there was no TET 1 expression in these three cell lines. Meanwhile, the other three cell lines (LS174T, SW480, and LOVO) showed a low level of DNA methylation correlated with TET1 expression. They found that DNA methylation of the TET1 gene is correlated with silencing of TET1. They also examined TET 1 expression in tumor and normal tissues from colorectal cancer cases. TET1 down-regulation was significantly detected in cancerous tissues compared to normal tissues.

CpG island methylator phenotype (CIMP), a phenotype with an extensive CpG island promoter methylation, has mostly been studied in colorectal cancer (Hughes et al., 2013). Ichimura et al. (2015) also examined the expression level of the TET family genes in CIMP-positive and CIMP-negative colorectal cancer tissues. TET 1 was significantly down-regulated in CIMP-positive colorectal cancer cases (23/55, 42%) compared to CIMP-negative cases (2/113, 2%).

Li et al. (2016) analyzed TET1 the DNA methylation and expression status in multiple cancer cell lines and primary tumor samples. For example, 27% (3/11) of primary colon carcinomas showed TET1 promoter methylation by methylation-specific PCR and methylation was associated with reduced mRNA expression by semi-quantitative reverse transcription-PCR. The results demonstrated that there is a direct link between promoter CpG methylation and TET1 inactivation. They concluded that the tumor-specific methylation of TET1 could serve as an epigenetic marker.

In one study, Hsu et al. (2012) indicated that TET1 is downregulated in prostate and breast cancer tissues and TET1 inhibits prostate and breast cancer invasion and metastasis. To

our knowledge, TET1 methylation has not been addressed in endometrial hyperplasia or ovarian cancer, so far.

6.4. The ARID1A gene

ARID1A (AT-rich interactive domain 1A) gene is a tumor suppressor gene located at Ch1P36.11 and encodes a DNA-binding protein called BAF250a. ARID1A is a subunit of the SWI/SNF adenosine triphosphate-dependent chromatin-remodeling complexes. SWI/SNF complexes play an essential role in tissue development, differentiation and tumor suppression (Reisman et al., 2009). Frequent mutations of ARID1A has been discovered in ovarian clear cell carcinoma (Jones et al., 2010). Somatic inactivating mutations of ARID1A and loss of its expression occur most frequently in endometrium-derived tumors, including ovarian clear cell carcinomas, ovarian and uterine endometrioid carcinomas (Mao & Shih., 2013).

Zhang et al. (2013) examined the correlation between ARID1A hypermethylation and mRNA expression in breast cancers. They analyzed promoter methylation of 38 pairs of breast invasive ductal carcinoma and their normal breast tissues using MeDIP-qPCR. Low mRNA expression was detected by real time RT-PCR in 22 out of 38 (57.9%) cancerous tissues. In this low expression group, 19 of 22 (86.4%) samples showed ARID1A promoter hypermethylation. They came to the conclusion that promoter hypermethylation is the main reason for low expression of ARID1A mRNA in invasive ductal carcinomas.

Aims of this study

As several studies have shown that there is a connection between aberrant DNA methylation status and human cancers, the first aim of this study was to design a custom-made MS-MLPA assay to investigate alterations in DNA methylation in endometrial and ovarian tumors. Four genes associated with ovarian and endometrial cancers in the literature were selected to be examined (HNF1B, TET1, L1CAM, and ARID1A). The study material consisted of sporadic ovarian cancer (79 cases), endometrial hyperplasias (59 cases), normal endometrium (18 cases) and normal fallopian tube (14 cases). The promoter methylation status of the selected genes were analyzed using a custom-designed MS-MLPA test.

Another purpose of this study was to evaluate if the custom-made MS-MLPA test would work properly on the clinical samples. Since the clinical samples are usually formalin-fixed, paraffin-embedded, they have poor quality. As currently the majority of methods for DNA methylation analysis are not suitable for paraffin-embedded tissues, MS-MLPA test could serve as powerful tool for DNA methylation profiling in paraffin-embedded tissues where small amounts of DNA are available.

Material and methods

1. The cell lines used in bisulfite sequencing studies

Fifteen cancer cell lines and five normal cell lines and healthy control DNA samples, listed in Table 1, were obtained from the American Type Culture Collection (Rockville, USA) before this study and used in this study for bisulfite sequencing.

Table 1. Cell lines used for bisulfite conversion studies.

Name of the cell lines	Tissue of origin & disease	Name of the cell lines	Tissue of origin & disease
ES2	Ovary, clear cell adenocarcinoma	HCT116	Colon, colorectal carcinoma
OV4	Ovary, serous adenocarcinoma	SW480	Colorectal adenocarcinoma
SKOV3	Ovary, serous adenocarcinoma	BT-549	Hypermethylator breast cancer cell
CAOV3	Ovary, serous adenocarcinoma	CAL-51	Breast cancer cell with MSI
AN3CA	Uterus/endometrium adenocarcinoma	ZR-75-1	Low methylator breast cancer cell
HCA7	High methylator colorectal carcinoma	MCF-12A	Normal breast epithelial cells
KM12	Colon, adenocarcinoma	TKF	Blood
RKO	High methylator colorectal carcinoma	SPR.198	Blood
T84	Colon, colorectal carcinoma	Normal uterus	Epithelial
HCT15	High methylator colorectal carcinoma	Normal colon	Epithelial

2. Tumor and normal samples used in MS-MLPA studies

As listed in Table 2, the study material consisted of sporadic ovarian tumors (n = 79, including 24 endometrioid, 35 clear cell, and 20 high-grade serous tumors) and precursor lesions of endometrial carcinoma (n = 59, including 19 sporadic complex atypical hyperplasia, 17 sporadic simplex hyperplasia, and 23 sporadic complex hyperplasia). Reference group included 18 normal endometrium and 14 normal fallopian tissues.

Table 2. Number of patient samples used in this study.

Sample name	Sample number
Precursor lesions of endometrial carcinoma	
Sporadic complex atypical hyperplasias	19
Sporadic simplex hyperplasia	17
Sporadic complex hyperplasia	23
Sporadic ovarian cancer	
Sporadic serous ovarian cancer	20
Sporadic endometrioid ovarian cancer	24
Sporadic clear cell ovarian cancer	35
Normal samples	
Normal endometrium tissues	18
Normal fallopian tissues	14

3. DNA extraction

Genomic DNA from cancer and normal cell lines and DNAs had been previously isolated using the method of Lahiri et al. (1991) before this study. DNA from paraffin-embedded tumor samples had also been previously isolated according to the modified protocol of Isola et al. (1994).

4. Bisulfite sequencing

Bisulfite sequencing is a method for determining DNA methylation. In this method, DNA is first treated with sodium bisulfite, during which the unmethylated cytosines are converted to uracils and 5-methylcytosines remain as cytosines. Then, converted DNA is amplified by PCR using bisulfite primers. After bisulfite conversion, the two strands of DNA are no longer complementary. Thus, both methylated and unmethylated sequences will be amplified. In the PCR reaction, 5-methylcytosines are amplified as cytosines and the uracils are amplified as thymines. Finally, the PCR product is sequenced to quantify the DNA methylation status (Clark et al., 2006).

4.1. Bisulfite conversion of DNA

Altogether, 600 ng of genomic DNA was used for bisulfite conversion of DNA, in which unmethylated cytosine is deaminated to produce uracil in DNA, and methylated cytosines are protected from this conversion and remain as cytosines. All the cell lines, including 15 cancer cell lines, three normal cell lines and two normal DNA samples, were bisulfite converted. To perform bisulfite treatment, the EZ DNA Methylation-Gold Kit was used, as described by the manufacturer (Zymo Research, Orange, CA, USA).

4.2. Bisulfite primer design and PCR of bisulfite-converted DNA

As the primer design is critical for successful bisulfite PCR, special attention was given to the following primer designing guidelines. The primers should design so that the target promoter region contains one or more HhaI recognition site which later are needed for the MS-MLPA probe design. Bisulfite PCR primers should design longer than normal PCR primers (26-30), and the amplicon size should be relatively short (190 bp to 217 bp). Preferably, primers should not contain CpG sites in their sequences. Usually, an annealing temperature between 55-60 °C works well. To identify the promoter regions for any CpG islands, the EMBOSS CpGplot program (www.ebi.ac.uk/Tools/seqstats/emboss_cpgplot/index.html) and the MethPrimer

(www.urogene.org/methprimer) were used. All the studied promoter regions located on the CpG islands, except L1CAMI. The L1CAM gene has two active promoter sites (Promoter 1 and 2). Promoter 1 was studied here. In the ARID1A gene, two CpG islands were studied.

Bisulfite converted DNA was amplified using the following protocol: initial denaturation at 95 °C for 5 minutes, 35 cycles of denaturation (30 seconds at 94 °C), annealing for 30 seconds (variable Tm for each gene), extension (30 seconds at 72 °C), and at the end, a final extension at 72 °C for 5 minutes. In the bisulfite PCR amplification, uracil is recognized as thymine and unmethylated cytosines are recognized as cytosines, allowing methylated and unmethylated cytosines to be distinguished. The reaction mix used for bisulfite PCR amplification and the bisulfite primer sequences are described in Table 3 and 4, respectively. The resulting PCR fragments were visualized by 2% agarose gel electrophoresis. After the gel separation, the PCR products were purified using ExoSAP-IT® PCR Clean-Up reagent (Usb®, USA, 78200), in which excess primers and dNTPs were eliminated from the PCR product. After purification, the PCR products were sequenced with an Applied Biosystems ABI 3730 Automatic DNA Sequencer.

Table 3. The reaction mix used in bisulfite-PCR reactions.

Constituent concentration	Manufacturer	Volume (µl)	Final concentration
10X PCR buffer(containing 15mM Mgcl2)	QIAGEN	2,5	1X
dNTP mixture	Finnzymes	0,5	200 µM
Forward primers (5 µM)	Sigma-Aldrich	2	0.4 µM
Reverse primers (5 µM)	Sigma-Aldrich	2	0.4 µM
HotStar Taq DNA polymerase (5 units/µl)	QIAGEN	0,25	1.25 units/reaction
Distilled water		Variable	-
Template of bisulfite converted DNA		Variable	0.05-0.15 µg/ 25 µl
Total volume		25	-

Table 4. Primers used in bisulfite-PCR reactions.

Gene	Primers sequences	Tm	Size	References
HNF1B	F: 5'- GGGGTYGAGTTYGATATTAAGT-3'	52	191 bp	Terasawa et al. 2006
	R: 5'- TACCTAAACATCCRATCCACCT-3'	52		Terasawa et al. 2006
TET1	F: 5'- AGGGTTGGTGTAGGTTTGGAGTT-3'	60	220 bp	This study
	R: 5'- AAAACRAACCCACCCCTAAAACAA-3'	60		This study
L1CAM, island I	F: 5'- TTAGAGAGTTGGAGGAAAATTTG-3'	58	261 bp	Schirmer et al. 2013
	R: 5'- ACACACACACACAAAACAAAAC-3'	58		Schirmer et al. 2013
L1CAM, island II	F: 5'- GTTTTAGGTTTTTGGGAGTATTTT- 3'	58	225 bp	This study
	R: 5'- CACCCTAACCCCTAATACCAAC- 3'	58		Schirmer et al. 2013
ARID1A, Island 4	F: 5'- GTTGTTAGGGGGTTAGGGTTG- 3'	60	297 bp	This study
	R: 5'- CCTTCCCTTCACAAAAAAAA- 3'	60		This study
ARID1A, Island 5	F: 5'- GTTTGGGGGGAATGAGTYGGGAG- 3'	64	217 bp	This study
	R: 5'- TCCCCCRAACTACCCTCCCCAC- 3'	64		This study

5. MS-MLPA

5.1. MS-MLPA probe design

MS-MLPA probes were designed for HNF1B, L1CAM, TET1, and ARID1A as described in the guidelines, *Designing synthetic MLPA probes* (MRC-Holland, the Netherland, version 10). Each designed probe contained at least one HhaI restriction site in the target recognition sequence (Figure 3). Custom-designed MS-MLPA probe sets are described in Table 5. CpG islands in the selected genes were identified using the EMBOSS CpGplot and MethPrimer programs. All of the designed MS-MLPA probes were located in the CpG islands of the analyzed genes, except L1CAMI. Structures of promoter CpG islands for the studied genes are depicted in Figures 4-7 below.

Each MS-MLPA probe consists of two separate oligonucleotides, including the left probe oligonucleotide (LPO) and the right probe oligonucleotide (RPO). Both oligonucleotides contain a universal primer sequence, hybridizing sequence, and an optional stuffer sequence that is used when needed to modulate the total probe length. Each MS-MLPA probe contains a recognition site for the methylation-sensitive restriction enzyme. For example, Hhal restriction enzyme recognizes GCGC sites.

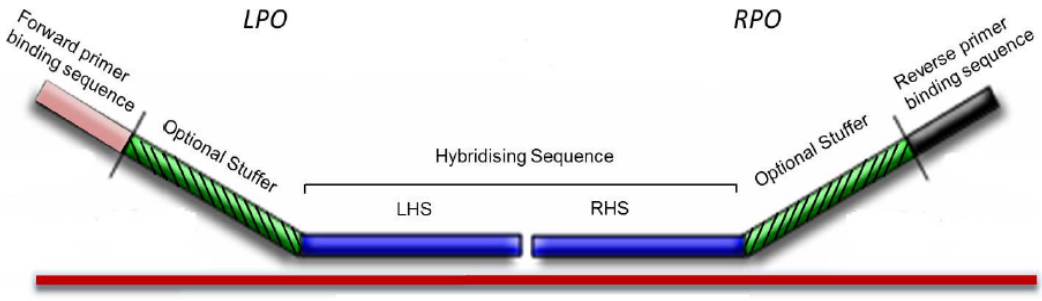


Figure 3. MS-MLPA probe components (source: Designing synthetic MLPA probes instructions, MRC-Holland, the Netherlands).

When designing MS-MLPA probes, specific attention was given to following guidelines. In each MS-MLPA reaction, there should be at least a total of 5 unique probes in a probemix to obtain reliable results, and there should not be more than 11 synthetic probes in a reaction. LPO and RPO should locate directly adjacent to each other, and the minimum length of each should be 21 nucleotides. Each designed probe should have a unique total amplicon length. The optimal probe length is between 100 and 140 nucleotides, as long probes will produce a lower signal and they will often have shoulder peaks. There should be a space of at least 4 nucleotides between two probes. The minimum Tm for each hybridizing sequence is 68 °C, and preferably Tm should be higher than 71 °C. Preferably, the secondary structure (ΔG) for each probe should be positive, although some probes with negative ΔG work. Details on probe sequences are shown in Table 6.

5.2. The MS-MLPA assay and data analysis

MS-MLPA was performed as described by the manufacturer (MRC-Holland, Amsterdam). MS-MLPA reagents were purchased from MRC-Holland, Amsterdam, and the Netherlands. Approximately 150 ng of genomic DNA was denatured for 10 min at 98 °C. SALSA MLPA buffer (1.5 µl), reference probes (1 µl), and designed probes (0.5 µl) were then added, and after incubation for 1 min at 95 °C , were allowed to hybridize to their respective target for about 16 h at 60 °C. After hybridization, the mixture was diluted at room temperature with H_2O and 3 µl ligase buffer A to a final volume of 20 µl and then equally divided between two tubes. In one tube, half of the sample was ligated using H_2O (8.25 µl), ligase buffer B (1.5 µl), and ligase-65 enzyme (0.25 µl), whereas in the second tube, ligation was combined with Hhal digestion replacing Hhal enzyme with H_2O. PCR was performed by the addition of 5 µl of the PCR mixture containing H_2O (3.75 µl), SALSA primer mix (1 µl), and SALSA polymerase (0.25) to the ligation mixture. PCR products were diluted in water and mixed with internal size standard (LIZ Sixe) and deionized formamide, and separated by capillary gel electrophoresis and quantified using Genemapper software (Applied Biosystems).

MS-MLPA data was analyzed in Microsoft Excel as described by the manufacturer's guideline (MRC-Holland, Amsterdam). To quantify the methylation status of each GCGC site the peak area of each probe is first divided by the combined value of the control probes within the samples generating relative peak values. Then, the relative peak values of the digested sample are divided by those of the corresponding undigested sample. This number is the methylation ratio, Dm (Nygren et al., 2005).

For a meaningful interpretation of the methylation data, the Dm values were used to determine two important thresholds, first, the technical threshold for methylation and second, the threshold for hypermethylation. The former describes the ability of the MS-MLPA method to detect methylation when compared to bisulfite sequencing (see p.44). Hypermethylation refers to increased methylation in tumor DNA compared to the corresponding normal DNA. The threshold for hypermethylation was defined for each gene individually and calculated as the average Dm plus one standard deviation (if this value was lower than the technical threshold, the technical threshold was used instead, see p. 50).

5.3. Statistical analysis

Statistical analysis were conducted using SPSS software, version 24.0 (IBM SPSS Inc. Chicago, IL, USA). Fisher's exact test was used to calculate two-sided *p* values for pairwise comparisons of hypermethylation frequencies between groups. *P* values < 0.05 were considered significant.

Table 5: Target sequence of probes used for the custom-made MS-MLPA test.

Gene	Location within CpG islands	Probe target sequence with HhaI site in bold	Distance of HhaI site from coding ATG (bp)*
HNF1B	Yes	(LHS)GACGACTATGACACACCTCCCATCCTCAAG- (RHS)GAGCTGCAG**GCGC**TCAACACCGAGGAG	+286
TET1	Yes	(LHS)CCGGGAGGCG**GCGC**TCG**GCGC**GGGCTGGAT- (RHS)GTGGCGGGCTCTGCGTCCTTGGCTCTCC	−11514 & −11507
L1CAM, island I	No	(LHS)CCTGGGCTCGCCTCCTACCCTGCCGCCCAC- (RHS)CTGGGCTCTGGGGT**GCGC**AGGAGCCGGTGC	−10927
L1CAM, island II	Yes	(LHS)GGGCTGCCCGGGGGCCCCCAGCCCAAG- (RHS)CC**GCGC**TCGGCTGGCACCAGGGGCTAGGGT	−10479
ARID1A, Island 4	Yes	(LHS)CGGGAGGGGACAGACCTGGATAGGGACG- (RHS)CCGGGAGGGAGGCG**GCGC**AGGCTCCAGA	−900
ARID1A, Island 5	Yes	(LHS)CCTACAGAGCCGGGAGCAGCTGAGCCGC- (RHS)CG**GCGC**CTCGGCCGCCGCCGCCGCCTCCT	−227

LHS: Left hybridizing sequence/ RHS: Right hybridizing sequence.
* Plus sign indicates an upstream location whereas a minus sign denotes a downstream location.

Table 6. Probe design for the custom-made MS-MLPA test (detailed characteristics).

Gene	GC %	Tm °C	ΔG kcal/mol	Probe sequences
HNF1B	53 %	76	1.88	LPO: 5' **GGGTTCCCTAAGGGTTGGA***CATCTTGAGTCCATCTTGAGTCCATCT*GACGACTATGACACACCTCCCATCCTCAAG 3'
	67 %	81.9	0.89	RPO: 5' GAGCTGCAG***GCGC***TCAACACCGAGGAG*TTCATGGTTCATGGTTCATGGTTCAT***TCTAGATTGGATCTTGCTGGCAC** 3'
TET1	68 %	97	– 1.01	LPO: 5' **GGGTTCCCTAAGGGTTGGA** *CATCTTGA*CCGGGAGGCG***GCGC***TCG***GCGC***GGGCTGGAT 3'
	58 %	91	0.98	RPO: 5' GTGGCGGGCTCTGCGTCCTTGGCTCTCC*TTCATG***TCTAGATTGGATCTTGCTGGCAC** 3'
L1CAM, island I	77 %	89.95	– 0.11	LPO: 5' **GGGTTCCCTAAGGGTTGGA***CATCTTGAGTCCA*CCTGGGCTCGCCTCCTACCCTGCCGCCCAC 3'
	77 %	90.84	– 0.31	RPO: 5' CTGGGCTCTGGGGT***GCGC***AGGAGCCGGTGC*TTCATGGTT***TCTAGATTGGATCTTGCTGGCAC** 3'
L1CAM, island II	85 %	93.96	– 1.82	LPO: 5' **GGGTTCCCTAAGGGTTGGA***CATCTTGAGTCCATCTTGAGTC*GGGCTGCCCGGGGGCCCCCAGCCCAAG 3'
	77 %	90.34	0.41	RPO: 5'CC***GCGC***TCGGCTGGCACCAGGGGCTAGGGT*TTCATGG TTCAT***TCTAGATTGGATCTTGCTGGCAC** 3'
ARID1A, Island 4	68 %	83.21	0.47	LPO: 5' **GGGTTCCCTAAGGGTTGGA***CATCTTGAGTCCATCTTGAGTCCAT*CGGGAGGGGACAGACCTGGATAGGGACG 3'
	79 %	90.75	0.69	RPO: 5' CCGGGAGGGAGGCG***GCGC***AGGCTCCAGA*TTCATGGTTCATGGTTCATG***TCTAGATTGGATCTTGCTGGCAC** 3'
ARID1A, Island 5	71 %	84.47	0.65	LPO: 5' **GGGTTCCCTAAGGGTTGGA***CATCTTGAGTC CATCTTGAGTC*CCTACAGAGCCGGGAGCAGCTGAGCCGC 3'
	90 %	96.63	– 1.65	RPO: 5' CG***GCGC***CTCGGCCGCCGCCGCCGCCTCCT*TTCATGGTTCATGGTTCA***TCTAGATTGGATCTTGCTGGCAC** 3'

LPO: Left probe oligo, RPO: Right probe oligo
PCR primer binding sequence (bold), stuffer (italic), hybridizing sequence (underlined), and HhaI site (bold italic)

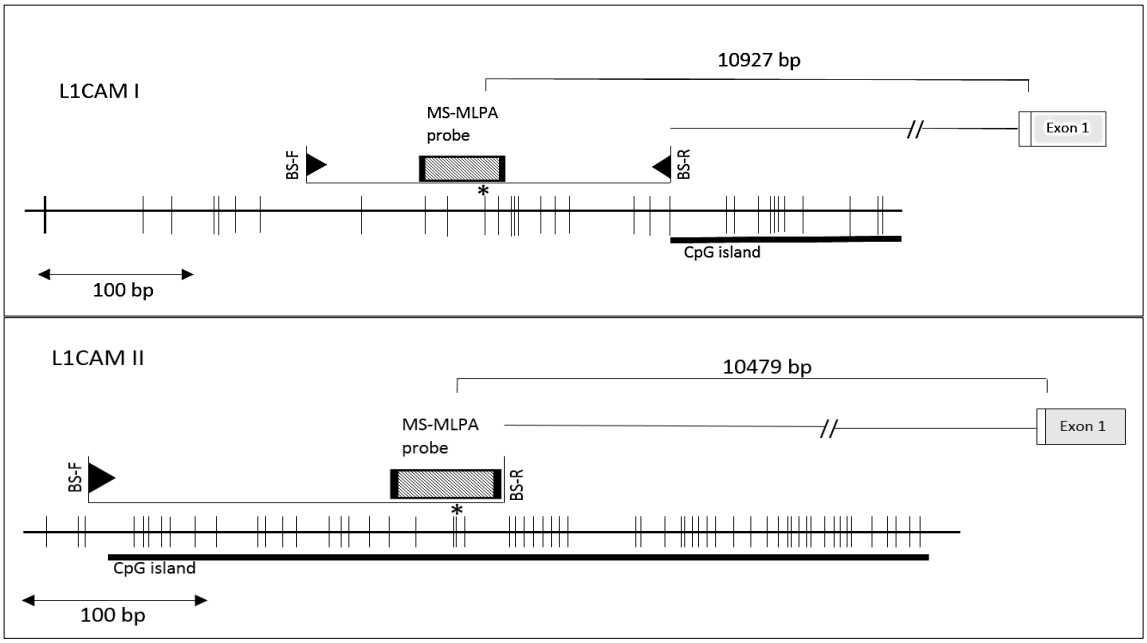


Figure 4: Structure of L1CAM promoter CpG island. CpG sites are shown as vertical lines. Bisulfite sequencing region (BS-F and BS-R) analyzed and MS-MLPA probe containing GCGC site (∗) are also indicated. The first coding exon is depicted as a separate box consisting of white box indicating untranslated region and the gray portion representing the translated region. Translation initiation site (ATG) is located at the start of the gray portion.

Reference for publication with the evidence for the importance of the chosen region for methylation analysis: Schirmer, U., Fiegl, H., Pfeifer, M., Zeimet, A., Müller-Holzner, E., Bode, P., . . . Altevogt, P. (2013). Epigenetic regulation of L1CAM in endometrial carcinoma: comparison to cancer–testis (CT-X) antigens. BMC Cancer, 13, 1-12).

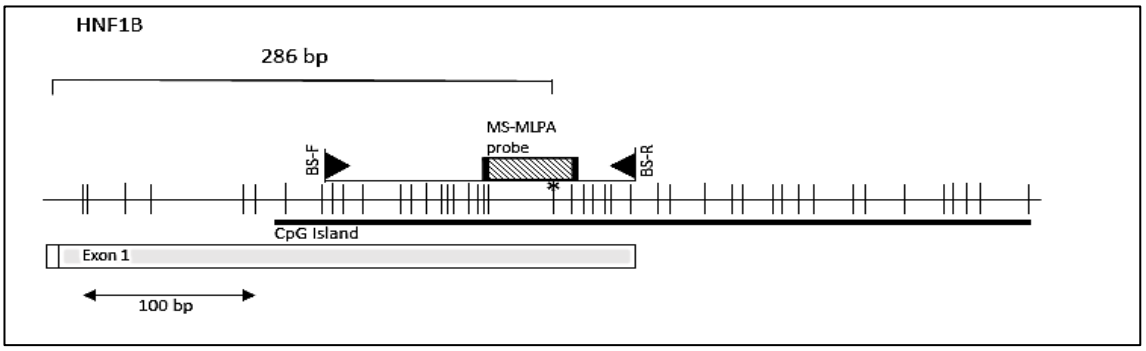


Figure 5: Structure of HNF1B promoter CpG island. CpG sites are shown as vertical lines. Bisulfite sequencing region (BS-F and BS-R) analyzed and MS-MLPA probe containing GCGC site (∗) are also indicated. The first coding exon is depicted as a separate box consisting of white box indicating untranslated region and the gray portion representing the translated region. Translation initiation site (ATG) is located at the start of the gray portion.

Reference for publication with the evidence for the importance of the chosen region for methylation analysis: Terasawa, K., Toyota, M., Sagae, S., Ogi, K., Suzuki, H., Sonoda, T., … Tokino, T. (2006). Epigenetic inactivation of TCF2 in ovarian cancer and various cancer cell lines. British Journal of Cancer, 94(6), 914–921.

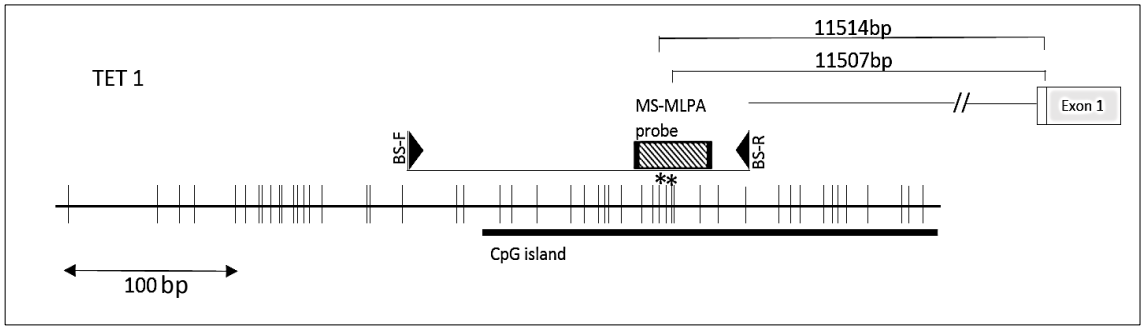


Figure 6: Structure of TET1 promoter CpG island. CpG sites are shown as vertical lines. Bisulfite sequencing region (BS-F and BS-R) analyzed and MS-MLPA probe containing GCGC sites (∗) are also indicated. The first coding exon is depicted as a separate box consisting of white box indicating untranslated region and the gray portion representing the translated region. Translation initiation site (ATG) is located at the start of the gray portion.

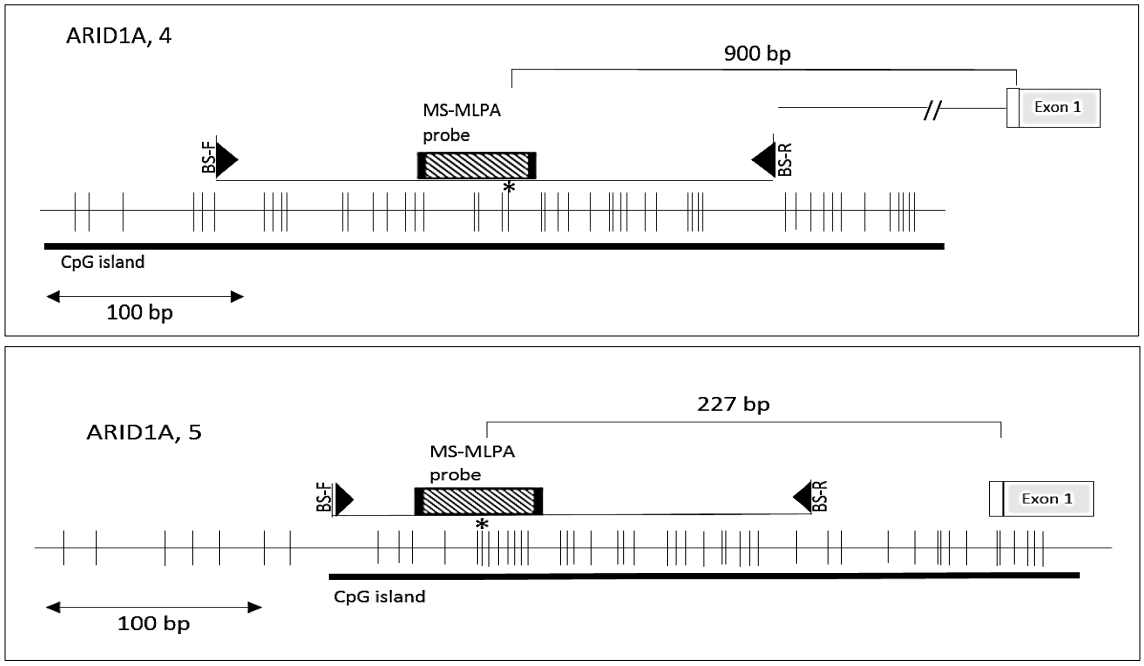


Figure 7: Structure of ARID1A promoter CpG island. CpG sites are shown as vertical lines. Bisulfite sequencing region (BS-F and BS-R) analyzed and MS-MLPA probe containing GCGC site (∗) are also indicated. The first coding exon is depicted as a separate box consisting of white box indicating untranslated region and the gray portion representing the translated region. Translation initiation site (ATG) is located at the start of the gray portion.

Results

1. Optimization of custom MS-MLPA test using cell lines

Six synthetic MS-MLPA probe pairs were designed for four genes (HNF1B, L1CAM CpG islands I and II, TET1, and ARID1A, CpG islands 4 & 5) according to the manufacturer's instructions (MRC-Holland, Amsterdam). DNA methylation patterns of 20 cell lines, including 15 cancer cell lines and 5 normal cell lines and DNAs, were analyzed by this custom-designed MS-MLPA test. For optimization of MS-MLPA test, DNAs of 20 cell lines were first bisulfite sequenced. After performing custom-made MS-MLPA test on these cell lines, the bisulfite sequencing results were compared with those of MS-MLPA test. After proving the correlation between bisulfite sequencing and MS-MLPA results, the test was applied to the clinical samples. A detailed description of results from each step will follow.

To validate the accuracy of the custom-designed MS-MLPA kit in DNA methylation detection, bisulfite sequencing was carried out. In each studied promoter region, the DNA methylation patterns of several CpG sites were analyzed by bisulfite sequencing. The CpG site containing the restriction site (GCGC) for the HhaI enzyme in the corresponding MS-MLPA test should ideally represent the DNA methylation pattern of the whole region. DNA methylation status of one colorectal cancer cell line (RKO) detected by bisulfite sequencing and MS-MLPA test is presented in Figure 8.

As displayed in Table 7, the bisulfite sequencing results from normal and cancer cell lines were compared with those of custom-made MS-MLPA test. The purpose of comparing bisulfite sequencing vs. MS-MLPA results of the cell lines was to determine the technical threshold value for a reliable detection of methylation. Based on data shown in Table 7, the technical threshold was set at 0.20. When the methylation dosage ratio was ≥ 0.20, the cell lines were considered to be methylated. The cell lines with a methylation dosage ratio lower than 0.20 were interpreted to be unmethylated. Twenty cell lines were used for DNA methylation analysis, including 4 ovarian cancer cell lines, 7 CRC cell lines, 3 breast cancer cell lines, and 5 normal cell lines and DNAs. In Table 7, the presence of C

and T bases in the bisulfite sequencing tracing are indicated as C and T, respectively. If a heterogenous pattern of both C and T bases were present in bisulfite sequencing, this is shown as C/T or T/C, where the base with the highest peak is written first.

As presented in Table 7, MS-MLPA results were in agreement with the bisulfite sequencing results. All cell lines that showed methylation (C indicates methylation) by bisulfite sequencing were also shown to be methylated by MS-MLPA with a variation of Dm between 0.28 (L1CAM gene, island II in AN3CA cell line) to 1.16 (L1CAM I gene, island I in HCA7cell line). As the cut-off limit was set at 0.20 in MS-MLPA test, all cell lines that were unmethylated (T indicates unmethylation) by bisulfite sequencing, also showed to be unmethylated by MS-MLPA. In unmethylated CpG sites, the methylation ratio varied between 0 (In the vast majority of cell lines in TET1 gene, ARID1A gene, island 4 & 5) to 0.18 (ARID1A gene, island 4 in KM12 cell lines).

If the studied CpG sites showed a heterogenous pattern of both C and T by bisulfite sequencing, the methylation dosage ratios by MS-MLPA varied between 0.20 (L1CAM gene, island II, normal uterus blood sample) to 0.69 (L1CAM gene, island I, SPR.198 blood sample; however the methylation ratios were not lower in the sites interpreted as T/C compared to C/T. Heterozygous sites (C/T & T/C) in bisulfite sequencing results were interpreted as methylated by MS-MLPA test.

The ARID1A gene (both island 4 and 5) was unmethylated in all cancer cell lines and normal samples. The TET1 gene was also unmethylated in most of the cancer cell lines and all normal samples. Three colorectal cancer cell lines, including KM12, RKO, and HCT15 were methylated in the TET1 gene. The HNF1B gene was methylated in the majority of the cancer cell lines and is shown to be unmethylated in all normal samples. The L1CAM gene island I was methylated in all cancer cell lines (except CAOV3) and normal samples, and the methylation dosage ratios by MS-MLPA were higher in cancer cell lines compared to normal samples. The L1CAM gene island II was methylated in the majority of the cancer cell lines and normal samples and the methylation dosage ratios did not differ much between methylated cancer cell lines and normal samples.

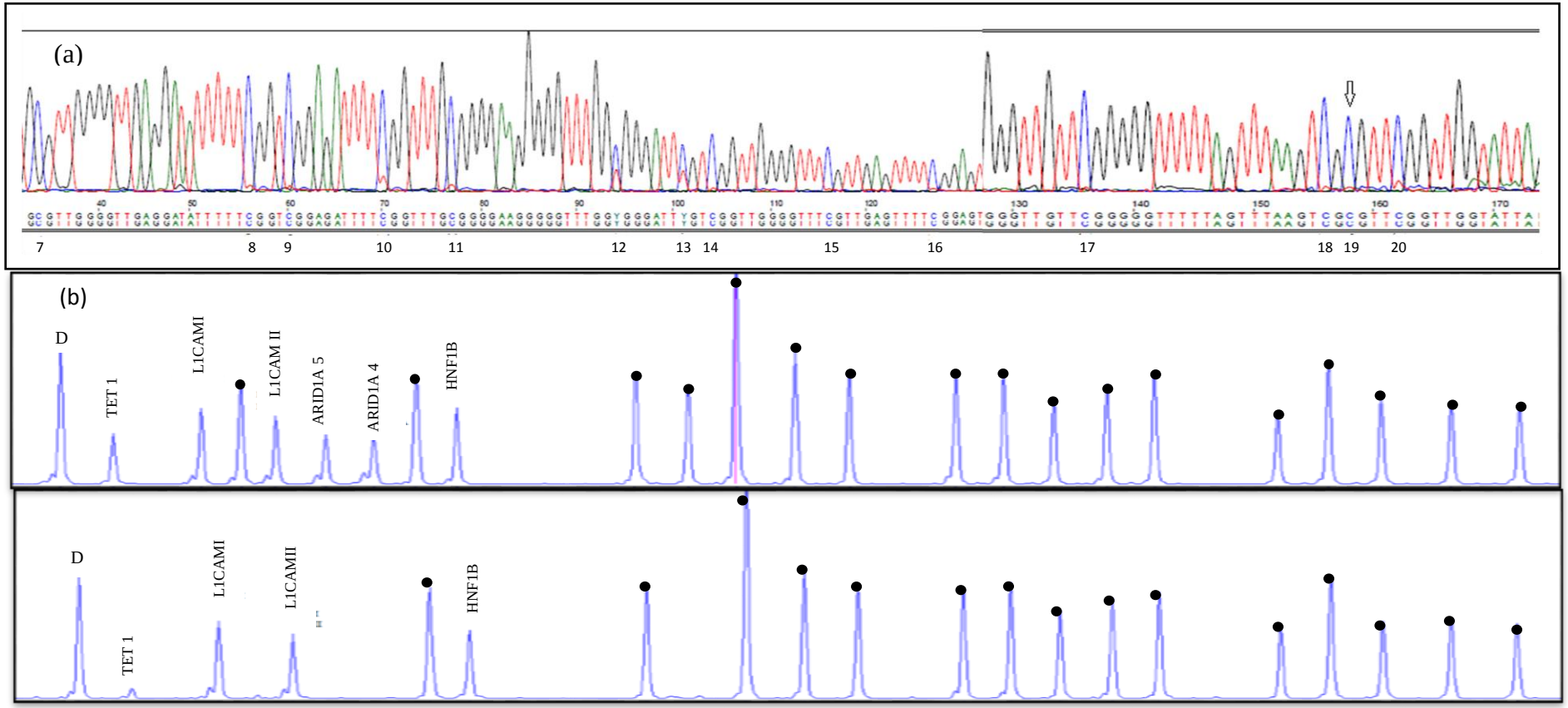


Figure 8. Bisulfite sequencing and MS-MLPA results of the colorectal cancer cell line RKO. (a) The CpG island studied by bisulfite sequencing in L1CAM island II contained 20 CpG sites, including one recognition site (GCGC) for the methylation-sensitive recognition enzyme HhaI (CpG sites are numbered and GCGC site is shown with arrow). As it can be seen, most of the CpG dinucleotides are methylated (C indicates methylation status), including the CpG site 19 analyzed later by MS-MLPA. **(b)** The top picture is the MS-MLPA results without HhaI treatment. The lower panel is after digestion with HhaI, resulting in peaks for control probes and methylated CpG sites. No peaks for unmethylated CpG sites are presented. (● : control peaks, D: fragment peaks).

Table 7. The comparison of methylation status acquired by bisulfite sequencing and MS-MLPA.

Cell line	Cell type	**HNF1B** Bisulfite sequencing (GCGCsite*)	**HNF1B** MS-MLPA (Dm)	**L1CAMI** Bisulfite sequencing (GCGC site)	**L1CAMI** MS-MLPA (Dm)	**L1CAMII** Bisulfite sequencing (GCGC site)	**L1CAMII** MS-MLPA (Dm)
ES2	Ovary, clear cell adenocarcinoma	C/T	0.43	C	1.02	C/T	0.17
OV4	Ovary, serous adenocarcinoma	C	0.73	C	1.03	C	0.41
SKOV3	Ovary, serous adenocarcinoma	T	0	C	1.08	C	0.38
CAOV3	Ovary, serous adenocarcinoma	C	0.92	T	0	T	0
AN3CA	Uterus/endometrium adenocarcinoma	C	0.76	C	1.09	C	0.28
HCA7	High methylator colorectal carcinoma	T	0	C	1.16	T	0.04
KM12	Colon, adenocarcinoma	C	0.88	C	0.97	C	0.30
RKO	High methylator colorectal carcinoma	C	0.91	C	1.00	C	0.91
T84	Colon, colorectal carcinoma	T	0	C	1.05	T	0
HCT15	High methylator colorectal carcinoma	T	0	C	1.03	C	1
HCT116	Colon, colorectal carcinoma	C	0.52	C	1.02	T	0
SW480	Colorectal adenocarcinoma	C	0.66	C	0.86	T	0.02
BT-549	Hypermethylator breast cancer cell	T/C	0.44	C	1.00	T	0.04
CAL-51	Breast cancer cell with MSI	T	0	C	0.95	C	0.42
ZR-75-1	Low methylator breast cancer cell	C	0.58	C	0.76	C/T	0.32
MCF-12A	Normal breast epithelial cells	C	0.29	C	0.92	C	0.44
TKF	Blood	T	0.03	C/T	0.54	C/T	0.24
SPR.198	Blood	T	0.05	C/T	0.69	T/C	0.24
Normal uterus	Epithelial	T	0	C/T	0.55	C/T	0.20
Normal colon	Epithelial	T	0	C/T	0.44	T	0

*In regard to bisulfite sequencing, the status of the CpG-dinucleotide-associated cytosine (GCGC) is indicated.

Cell line	Cell type	**TET1** Bisulfite sequencing (two *GCGCsites)	**TET1** MS-MLPA (Dm)	**ARID1A 5** Bisulfite sequencing (GCGC site)	**ARID1A 5** MS-MLPA (Dm)	**ARID1A 4** Bisulfite sequencing (GCGC site)	**ARID1A 4** MS-MLPA (Dm)
ES2	Ovary, clear cell adenocarcinoma	T&T	0	T	0	T	0
OV4	Ovary, serous adenocarcinoma	T&T	0	T	0	T	0
SKOV3	Ovary, serous adenocarcinoma	T&T	0	T	0	T	0
CAOV3	Ovary, serous adenocarcinoma	T&T	0	T	0	T	0
AN3CA	Uterus/endometrium adenocarcinoma	T&T	0	T	0	T	0
HCA7	High methylator colorectal carcinoma	T&T	0	T	0	T	0
KM12	Colon, adenocarcinoma	C&C	0.49	T	0.08	T	0.18
RKO	High methylator colorectal carcinoma	C&C	0.38	T	0	T	0
T84	Colon, colorectal carcinoma	T&T	0	T	0	T	0
HCT15	High methylator colorectal carcinoma	C&C	0.36	T	0	T	0
HCT116	Colon, colorectal carcinoma	T&T	0	T	0	T	0
SW480	Colorectal adenocarcinoma	T&T	0	T	0	T	0
BT-549	Hypermethylator breast cancer cell	T&T	0	T	0	T	0
CAL-51	Breast cancer cell with MSI	T&T	0	T	0	T	0
ZR-75-1	Low methylator breast cancer cell	T&T	0	T	0	T	0
MCF-12A	Normal breast epithelial cells	T&T	0	T	0	T	0
TKF	Blood	T&T	0	T	0	T	0
SPR.198	Blood	T&T	0	T	0	T	0
Normal uterus	Epithelial	T&T	0	T	0	T	0
Normal colon	Epithelial	T&T	0	T	0.04	T	0

*In regard to bisulfite sequencing, the status of the CpG-dinucleotide-associated cytosine (GCGC) is indicated.

2. Hypermethylation in precursor lesions of endometrial carcinoma

Methylation in the HNF1B, L1CAM (island I & II), ARID1A (island 4 &5), and TET1 genes was studied in 79 precursor lesions of endometrial carcinoma (23 sporadic complex hyperplasia, 19 sporadic complex atypical hyperplasia, 17 sporadic simplex hyperplasia) and 18 normal endometrium tissues using MS-MLPA test. The cut-off Dm values for hypermethylation relative to normal endometrium were determined as described in Material and methods (p.39). Gene-specific thresholds and their derivation are shown in Table 8. The HNF1B, TET1 and ARID1A genes showed low mean methylation dosage ratios in the normal endometrium tissues and the technical threshold (0.20) also served as the threshold for hyper-methylation. For L1CAM gene, the cut-off level of hypermethylation was (0.59 for island I and 0.39 for island II) calculated using the mean methylation dosage ratio, Dm, from all normal tissues + 1 standard deviation.

Table 8. Gene-specific thresholds for hypermethylation for samples with normal endometrium as the reference tissue (endometrial hyperplasias, and endometrioid and clear cell ovarian cancers).

Normal endometrium	HNF1B	TET1	L1CAM I	L1CAM II	ARID1A 4	ARID1A 5
01- Secretory	0.07	0	0.73	0.35	0.03	0.09
02- Secretory	0.06	0	0.56	0.34	0	0
03- Secretory	0.05	0	0.49	0.41	0.05	0
04- Secretory	0.10	0	0.51	0.3	0.04	0.05
05- Secretory	0.13	0	0.42	0.37	0.07	0
06- Secretory	0.08	0	0.44	0.31	0.06	0
07- Secretory	0.03	0	0.39	0.25	0.02	0.05
08- Secretory	0.05	0	0.48	0.3	0	0
09- Secretory	0.08	0	0.48	0.32	0	0
10- Proliferatory	0.00	0	0.56	0.33	0	0.18
11- Proliferatory	0.07	0	0.56	0.36	0	0
12- Proliferatory	0.08	0	0.54	0.29	0	0
13- Proliferatory	0.09	0	0.47	0.29	0.05	0
14- Proliferatory	0.10	0	0.55	0.47	0.06	0.05
15- Proliferatory	0	0	0.44	0.26	0.04	0
16- Proliferatory	0.05	0	0.33	0.33	0	0
17- Proliferatory	0.02	0	0.52	0.18	0	0
18- Proliferatory	0	0	0.20	0.12	0	0
Mean methylation dosage ($\bar{x}$)	0.06	0	0.48	0.31	0.02	0.02
SD	0.04	0	0.11	0.08	0.03	0.05
$\bar{x}$+1SD	0.10	0	0.59	0.39	0.05	0.07
Cut-off value for hypermethylation	0.20*	0.20*	0.59	0.39	0.20*	0.20*

* Represents the technical threshold (Material and methods, p.39)

Methylation results of endometrial hyperplasias are shown in two alternative ways below, as average Dm values (based on absolute Dm values given by MS-MLPA, Fig.9A) and as percentages of samples with hypermethylation (only samples with Dm values equal to or higher than cut-off values specified in Table 8 were counted as hypermethylated, Fig.9B). Figure 9B also gives results from statistical testing. In TET1 and ARID1A DNA methylation was not practically observed at all in either normal endometrium or hyperplasia cases. HNF1B showed slightly elevated methylation in hyperplasias. For L1CAM, there was an increasing trend for methylation (Fig. 9A) and especially hypermethylation (Fig. 9B) from normal endometrium to simplex hyperplasia to complex hyperplasia without atypia to complex hyperplasia with atypia. Compared to normal endometrium, all types of endometrial hyperplasia showed significantly higher frequencies of hypermethylation for L1CAM I (Fig. 9B). The same was true with complex hyperplasia with atypia for L1CAM II. The remaining comparisons indicate no formal statistical significance.

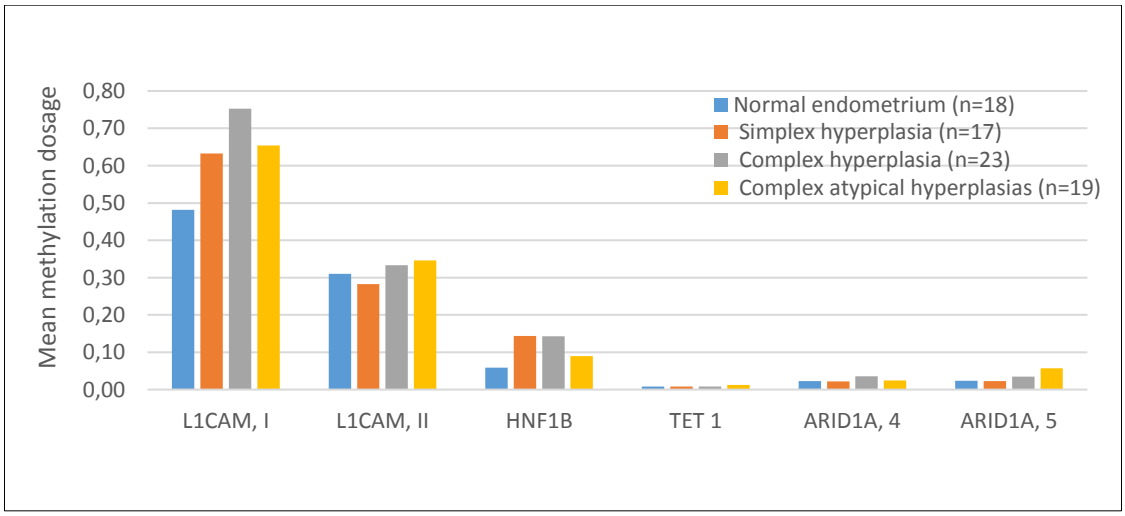
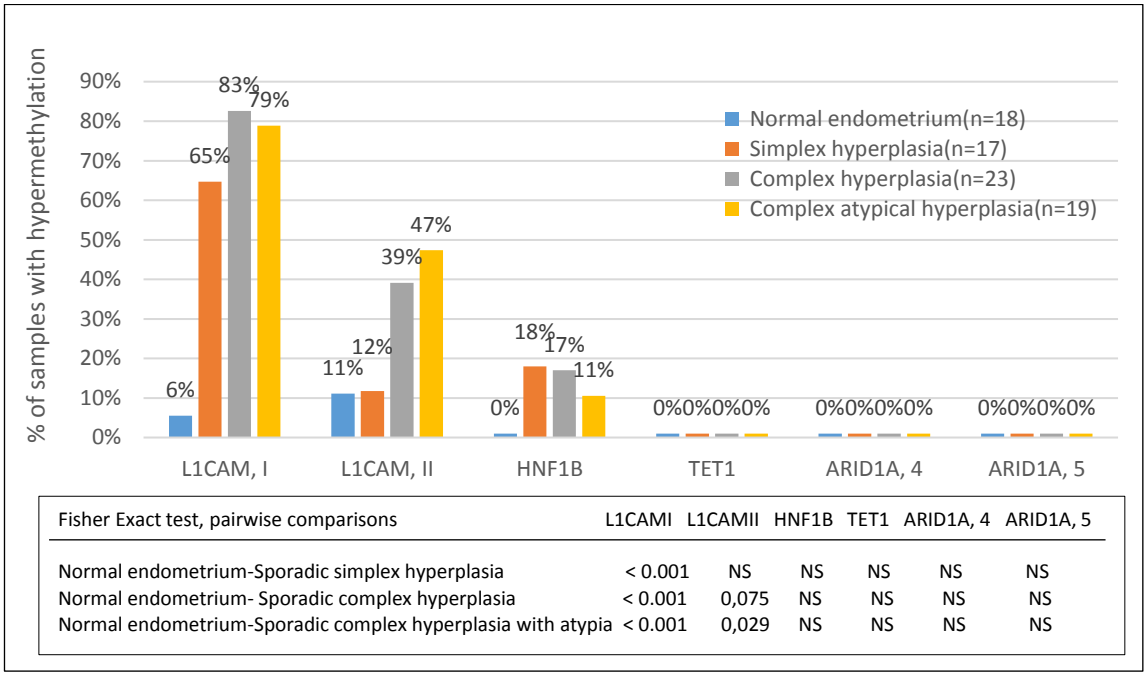


Figure 9A. The average Dm values for endometrial hyperplasias and normal endometrium shown separately for each gene.



Fisher Exact test, pairwise comparisons	L1CAMI	L1CAMII	HNF1B	TET1	ARID1A, 4	ARID1A, 5
Normal endometrium-Sporadic simplex hyperplasia	< 0.001	NS	NS	NS	NS	NS
Normal endometrium- Sporadic complex hyperplasia	< 0.001	0,075	NS	NS	NS	NS
Normal endometrium-Sporadic complex hyperplasia with atypia	< 0.001	0,029	NS	NS	NS	NS

Figure 9B. Frequencies of hypermethylation for endometrial hyperplasias and normal endometrium using gene-specific cut-off indicated in Table 8. Pairwise comparisons were calculated by Fisher Exact test (two-sided *P* values). *P* values < 0.05 were considered significant. NS= Not statistically significant.

3. Hypermethylation in clinical sporadic ovarian carcinoma samples

MS-MLPA was used to profile DNA methylation status of HNF1B, L1CAM (island I & II), ARID1A (island 4 &5), and TET1 gene in sporadic ovarian carcinomas. The clinical samples comprised sporadic ovarian cancer of the clear cell (n = 35), endometrioid (n = 24), and serous types (n = 20). It has been proposed that non-serous tumors (endometrioid and clear cell tumors) arise from endometrium and serous tumors arise from the fallopian tube, which was taken into account in the calculation of the threshold for hypermethylation. To have an optimal discrimination between normal and tumor tissues, the hypermethylation thresholds were calculated using the mean Dm in normal endometrium (for non-serous samples, Table 8 above) and normal fallopian tubes (for serous tumors, Table 9) plus 1 standard deviation. For genes with value lower than the technical threshold (0.20), the latter was used as the threshold for hypermethylation (similar to endometrial lesions described in the previous section).

Table9. Gene-specific thresholds for hypermethylation for samples with normal fallopian tubes as the reference tissue (serous ovarian cancer).

Normal fallopian tube	HNF1B	TET1	L1CAM I	L1CAM II	ARID1A 4	ARID1A 5
01-Normal fallopian tube	0.09	0	0.3	0.26	0	0
02-Normal fallopian tube	0.15	0	0.37	0.23	0	0
03-Normal fallopian tube	0.08	0	0.39	0.2	0	0
04-Normal fallopian tube	0.09	0	0.39	0.3	0	0.05
05-Normal fallopian tube	0.15	0	0.64	0.16	0	0
06-Normal fallopian tube	0.18	0	0.54	0.18	0	0
07-Normal fallopian tube	0.23	0	0.49	0.1	0	0
08-Normal fallopian tube	0.06	0	0.3	0.3	0	0
09-Normal fallopian tube	0.07	0	0.28	0.2	0	0
10-Normal fallopian tube	0.13	0.07	0.45	0.4	0.05	0
11-Normal fallopian tube	0.07	0	0.4	0.28	0	0
12-Normal fallopian tube	0.12	0	0.48	0.33	0	0
13-Normal fallopian tube	0.11	0	0.4	0.38	0.05	0.04
14-Normal fallopian tube	0.09	0.17	0.54	0.34	0	0.06
Mean methylation dosage ($\bar{x}$)	0.12	0.02	0.43	0.26	0.01	0.01
SD	0.05	0.05	0.10	0.08	0.02	0.02
$\bar{x}$+1SD	0.17	0.07	0.53	0.34	0.03	0.03
Cut-off value for hypermethylation	0.20*	0.20*	0.53	0.34	0.20*	0.20*

* Represents the technical threshold (Material and methods, p.39).

DNA methylation profiles for non-serous tumors are shown in Figures 10A and B. TET1 and ARID1A showed little methylation in normal endometrium and tumor tissues; among these genes, clear cell carcinoma with a 22% rate of hypermethylation for TET1 provided the strongest evidence of altered methylation (Fig. 10B). Increased methylation was more common for L1CAM and HNF1B (Figs. 10A & 10B). In HNF1B and L1CAM II an increasing trend for hypermethylation was observed from normal endometrium to clear cell to endometrioid tumors (Fig. 10B).

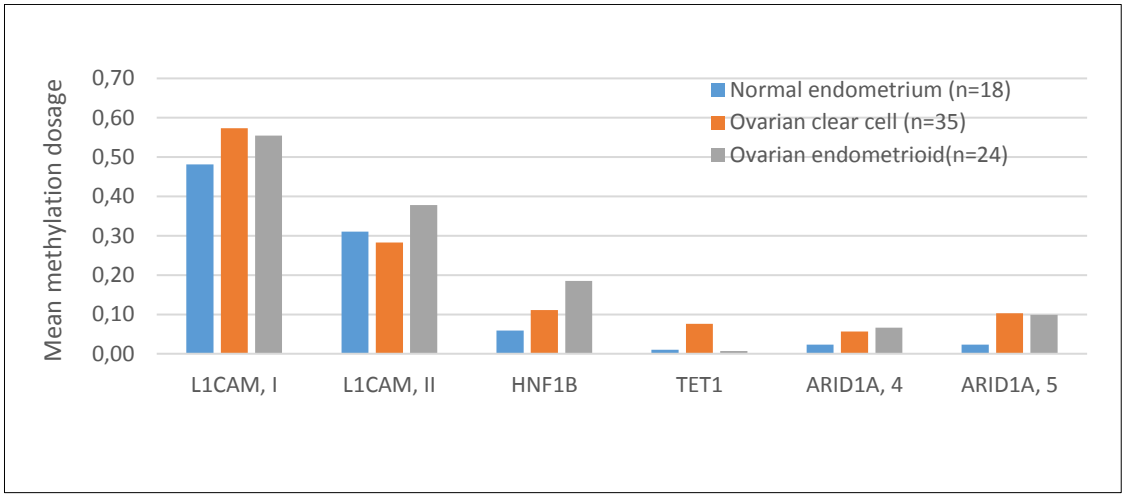
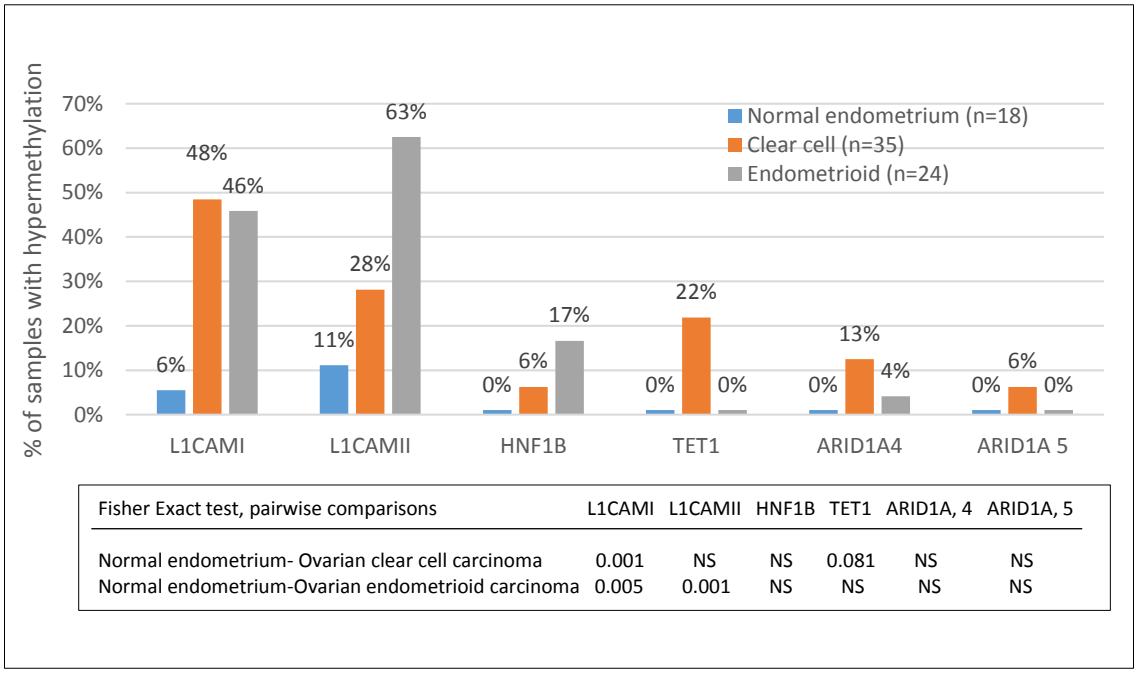


Figure 10A. The average of Dm values in sporadic non-serous ovarian tumors and normal endometrium.



Fisher Exact test, pairwise comparisons	L1CAMI	L1CAMII	HNF1B	TET1	ARID1A, 4	ARID1A, 5
Normal endometrium- Ovarian clear cell carcinoma	0.001	NS	NS	0.081	NS	NS
Normal endometrium-Ovarian endometrioid carcinoma	0.005	0.001	NS	NS	NS	NS

Figure 10B. Frequencies of hypermethylation for sporadic non-serous ovarian tumors and normal endometrium using gene-specific cut-off indicated in Table 8. Pairwise comparison were calculated by Fisher Exact test (Two-sided P values). *P* values < 0.05 were considered significant. NS= Not statistically significant.

DNA methylation results for serous tumors are displayed in Figures 11A and 11B. For TET1 and ARID1A, methylation was relatively rare in normal fallopian tubes and serous tumors (Fig. 11A & 11B). HNF1B and L1CAM II were frequently hypermethylated (25% & 35%, respectively) in serous samples compared to normal fallopian tubes (7% &14%, respectively, Fig. 11B). For L1CAMI, the frequency of hypermethylation in normal fallopian tubes exceeded that in serous tumors (Fig. 11B). Compared to normal fallopian tubes, serous tumors were not associated with significantly higher frequencies of hypermethylation for any studied genes.

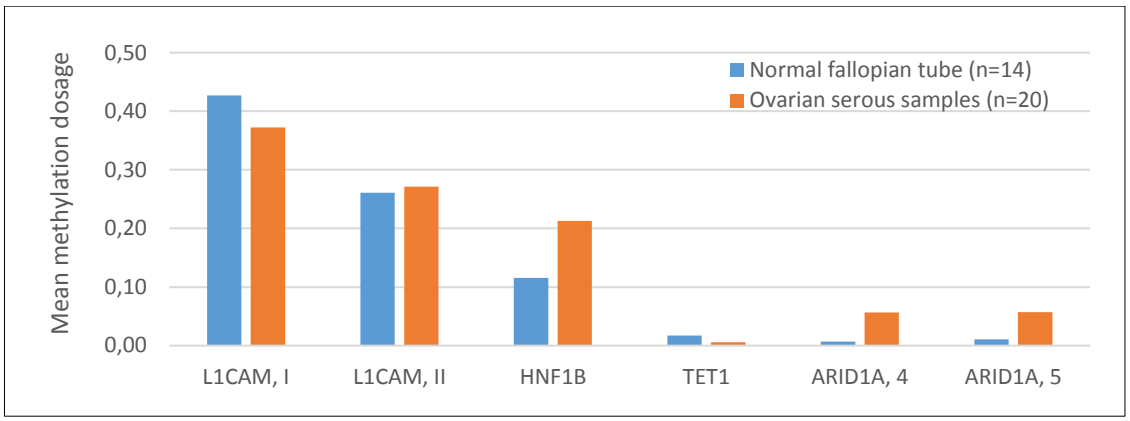


Figure 11A. The average of Dm values in sporadic serous ovarian tumors and normal fallopian tubes.

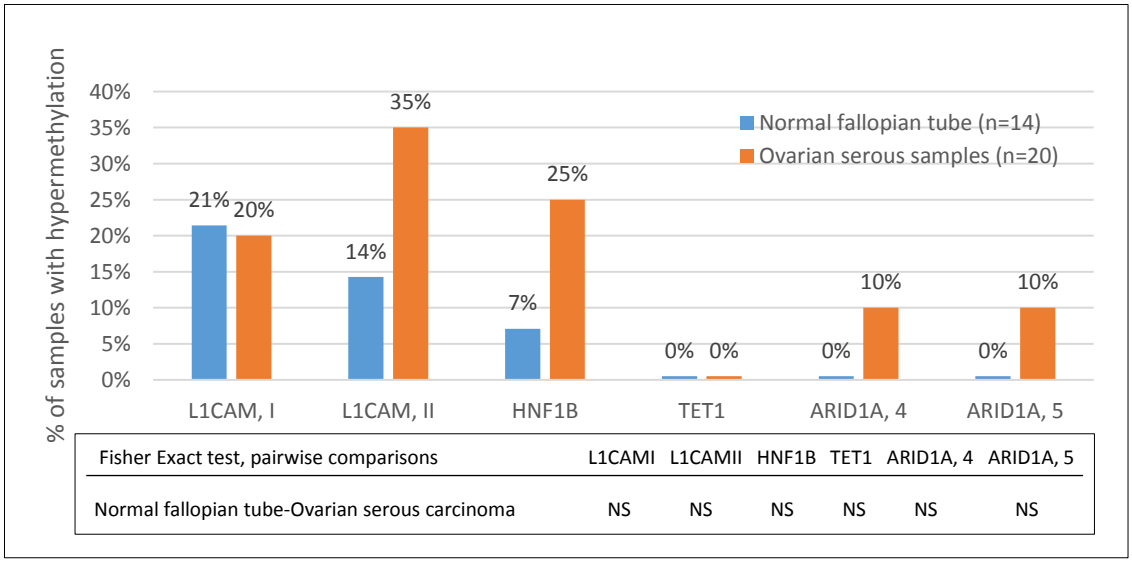


Figure 11B. Frequencies of hypermethylation in sporadic serous ovarian tumors and normal fallopian tubes using gene-specific cut-off indicated in Table 9. Pairwise comparison were calculated by Fisher Exact test (Two-sided P values). *P* values < 0.05 were considered significant. NS= Not statistically significant.

Discussion

1. Methylation status of the selected genes in cell lines

Our collection included four ovarian cancer cell lines, three of serous and one of clear cell origin (Table 1). HNF1B was completely methylated in 2/3 serous ovarian cancers and regarding the single unmethylated one (SKOV3), serous origin has been questioned (Niskakoski et al., 2014). The clear cell carcinoma (ES-2) showed partial methylation. HNF1B was completely unmethylated in normal sample DNAs. HNF1B was also completely methylated in 57% (4/7) of CRC cell lines, and completely unmethylated in normal cells. Our findings for the methylation status of HNF1B in cancer cell lines is in line with Terasawa et al. findings in 2006. They studied the methylation status of HNF1B gene in the SKOV3, CAOV3, RKO, HCT116, SW480 cell lines, and normal (colon and ovarian) samples using COBRA and bisulfite sequencing. HNF1B was methylated in the ovarian cancer cell lines in a histology specific pattern, resulting in loss of the gene's expression (observed in the same region as in this study).

The methylation level of L1CAM was higher in ovarian, colon, and breast cancer cell lines compared to normal samples. Kato et al, (2009) reported that L1CAM is methylated in 50% of CRC cell lines (2 of 4) and in 43.7% of primary CRCs (31 of 71 cases), which is in agreement with our results. In our study, of 7 CRCs cell lines, 3 (57.14%) were methylated.

ARID1A (Island 4 and 5) was unmethylated in both cancer cell lines and normal samples. We are not aware of any published methylation data for ARID1A regarding the same cell lines we studied. TET1 was methylated in 43% (3/7) of CRC cell lines but completely unmethylated in normal samples. Our findings of TET1 methylation are in line with Ichimura et al. (2015) and Li et al. (2016) results. They studied the DNA methylation status of SW480, RKO, HCT116, HCT15, T84, MCF-7, ZR-75-1 cell lines using pyrosequencing and MSP. The TET1 was methylated in RKO and HCT15 whereas it showed to be unmethylated in the rest of cancer cell lines and normal cell lines and DNAs.

2. Hypermethylation in endometrial lesions

We performed the DNA methylation analysis of the selected genes in lesions of the endometrium (simplex hyperplasia, complex hyperplasia without atypia and complex hyperplasia with atypia) using MS-MLPA approach. Normal endometrium tissues were used as a reference group. In endometrial hyperplasia series, methylation of L1CAM was increasingly hypermethylated along with histological abnormality (from simplex hyperplasia to complex hyperplasia to complex atypical hyperplasia, Fig.9B). This result is in line with the finding of Nieminen et al. (2009) and Joensuu et al. (2015) where the promoter DNA methylation of TSGs and the methyltransferase protein expression, respectively, increased with histological abnormality (observed in the same endometrial hyperplasia samples as in this study).

The increasing trend of methylation along with histological abnormality was observed in L1CAM gene in both island I and II (in L1CAMI, the frequency of hypermethylation in complex atypical hyperplasia is 4 % lower than that of in complex hyperplasia). Schirmer et al. in 2013 reported that L1CAM protein was absent in the vast majority of endometrioid endometrial carcinoma. It is proposed that endometrial hyperplasia can develop to endometrioid endometrial carcinoma. Since L1CAM expression has been shown to be inversely correlated with methylation of promoter I (Schirmer et al., 2013), our observation of hypermethylation of L1CAM in endometrial hyperplasia is in line with the reported absence of its expression in endometrioid endometrial carcinoma.

In our study, HNF1B showed low frequencies of hypermethylation in endometrial hyperplasias. Nemejcova et al (2016) studied the protein expression status of HNF1B in hyperplasia with atypia and hyperplasia without atypia. They detected expression of HNF1B in 88% of hyperplasias with atypias (14/16 cases), and 91% of hyperplasias without atypias (10/11cases). As it has been identified that the expression of HNF1B is inversely controlled by DNA methylation, our findings in HNF1B are in close concordance with this study. In TET1 and ARID1A, no methylation was detected in any histotypes of endometrial lesions.

3. Hypermethylation in sporadic ovarian carcinomas

We examined promoter methylation by the MS-MLPA method in 79 sporadic ovarian carcinomas of different histological types, using normal (unrelated) endometrium and normal (unrelated) fallopian tube tissues, as references.

Terasawa et al. (2006) reported that HNF1B was methylated in 41% of serous samples, 25% of mucinous tumors, and 29% of endometrioid tumors. There was no methylation in clear cell samples and normal ovarian endometrium tissues. In our study, the frequency of samples with hypermethylation increased from normal endometrium (0%) to normal fallopian tubes (7%) to clear cell (6%) to endometrioid (17%) and to serous samples (25%), which is in agreement with the report by Terasawa et al. in 2006. Since an inverse correlation of DNA methylation and protein expression has been identified in HNF1B (Terasawa et al. 2006) and HNF1B is shown to have different patterns of methylation between CCCs and non-CCCs, this gene can be used as a good marker of clear cell phenotype.

L1CAMI revealed low frequencies of hypermethylation in serous (20%) compared to endometrioid (46%) and clear cell ovarian carcinoma (58%). In terms of methylation dosage ratios, L1CAMI showed reduced methylation in serous ovarian carcinoma vs. normal fallopian tube (average Dm was 0.37 vs.0.43, respectively). Importantly, the average Dm in serous ovarian carcinoma was clearly lower than in endometrioid (0.55) and clear cell (0.58) ovarian carcinomas. These findings are in agreement with observation of high L1CAMI expression being characteristic of high-grade serous ovarian carcinoma (Bondong et al., 2012), under the assumption that expression and methylation are inversely correlated (Schirmer et al., 2013). DNA hypermethylation was not observed in TET1 at all, except in clear cell samples (22%).

4. Methodological aspects

DNA methylation status can be studied by a variety of different methods. Most of the common methods are based on bisulfite treated DNA, are time and work consuming, and often require good quality DNA. We chose the MS-MLPA method to analyze methylation mainly because it turned out to be suitable for paraffin embedded tissues. MS-MLPA is based on a methylation-specific endonuclease, does not use bisulfite converted DNA, and requires only 20ng of DNA. MS-MLPA approach was easy to learn and easy to perform. The custom-made MS-MLPA analysis can be considered to be a reliable approach to evaluate DNA methylation status in ovarian and endometrial carcinomas because there was a nearly complete agreement between bisulfite sequencing results and MS-MLPA results.

5. Conclusions and future prospects

In this research we analyzed the promoter DNA methylation status of four genes (HNF1B, L1CAM, TET1, and ARID1A) in ovarian carcinoma and precursor lesions of endometrial carcinoma. We found that HNF1B and L1CAM showed differential promoter methylation depending on the histological subtype of cancer. TET1 was methylated in clear cell carcinomas only. L1CAM showed an increasing trend in hypermethylation in endometrial precursor lesions (from normal endometrium to simple hyperplasia to complex hyperplasia without atypia to complex hyperplasia with atypia). Overall, our findings in this research are in line with existing literature, when available, and provide new information of the role of altered DNA methylation in endometrial and ovarian tumorigenesis.

This study focused on cell lines and sporadic samples of endometrial hyperplasias and ovarian cancer. One remaining task of interest is to study differences and similarities in DNA methylation, when ovarian cancers from the Lynch syndrome families are

compared with sporadic ovarian cancers. In addition, we would be interested to assess the DNA methylation status and other molecular alterations, when the precursor lesions of endometrial cancer with MMR gene mutation are compared with the sporadic series of endometrial specimens studied here. Lynch syndrome offers the advantage to examine consecutive endometrial specimens (because MMR mutation carriers participate in a regular surveillance by endometrial biopsies) as well as multiple cancers (endometrial and ovarian) developing in a single individual. Therefore, the custom-designed MS-MLPA test will be used in the future to analyze DNA methylation patterns in Lynch syndrome-associated ovarian and endometrial cancer.

Acknowledgement

This study was carried out at the Department of Medical and clinical Genetics, Biomedicum, University of Helsinki. Firstly, I am heartily thankful to my supervisor, Professor Päivi Peltomäki, whose guidance and support enabled me to develop an understanding of the subject. Thank you for this great opportunity to work in cancer research group with such a supportive atmosphere. My sincere thanks also goes to my other supervisor M.Sc. Anni Niskakoski for her introduction to the subject and advice in the laboratory and writing of this thesis. Besides my supervisors, I offer my special regards to Satu Valo who was always available to answer my questions. Her guidance helped me in all the time of research. I would like to thank Saila Saarinen for her tremendous technical help in the lab work. I also offer my regards to Taina Nieminen, Noora Porkka, and Alisa Olkinuora who supported me during the completion of the project. Last but not the least, I owe my deepest gratitude to my parents and siblings who have always encouraged me to study and made my studying possible.

References

Aarnio, M., Sankila, R., Pukkala, E., Salovaara, R., Aaltonen, L. A., Chapelle, A. D., Peltomäki, P., Mecklin, JP., & Järvinen, H. J. (1999). Cancer risk in mutation carriers of DNA-mismatch-repair genes. *International Journal of Cancer Int. J. Cancer, 81*(2), 214-218.

Baylin, S., & Jones, P. (2011). A decade of exploring the cancer epigenome — biological and translational implications. *Nature Reviews Cancer*, 11, 726-734.

Bogliolo, S., Cassani, C., Dominoni, M., Musacchi, V., Venturini, P., Spinillo, A., . . . Gardella, (2016). Veliparib for the treatment of ovarian cancer. *Expert Opinion on Investigational Drugs*, 25, 367-374.

Boland, C. R., & Lynch, H. T. (2013). The History of Lynch Syndrome. *Familial Cancer, 12*(2), 145-157.

Bonadona, V., Bonaiti, B., Olschwang, S., Grandjouan, S., Huiart, L., Longy, M., . . . Bonaiti-Pellie, C. (2011). Cancer risks associated with germline mutations in MLH1, MSH2, and MSH6 genes in Lynch syndrome. *Jama*, 305(22), 2304-2310.

Bondong, S., Kiefel, H., Hielscher, T., Zeimet, A. G., Zeillinger, R., Pils, D., . . . Altevogt, P. (2012). Prognostic significance of L1CAM in ovarian cancer and its role in constitutive NF- B activation. *Annals of Oncology, 23*(7), 1795-1802.

Burke, W. M., Orr, J., Leitao, M., Salom, E., Gehrig, P., Olawaiye, A. B., . . . Abu Shahin, F. (2014). Endometrial cancer: a review and current management strategies: part I. *Gynecologic Oncology*, 134(2), 385-392.

Clark, S.;Statham, A.;Stirzaker, C.;Molloy, P. L.;& Frommer, M. (2006). DNA methylation: Bisulphite modification and analysis. *Nature Protocols*, 1, 2353 - 2364.

Cuff, J., Salari, K., Clarke, N., Esheba, GE., Forster, AD., Huang, S., West, B., Higgins, J., Longacre, T., Pollack, J. (2013) Integrative Bioinformatics Links HNF1B with Clear Cell Carcinoma and Tumor-Associated Thrombosis. *PLoS ONE* 8(9): e74562.

Das, P. M. (2004). DNA Methylation and Cancer. *Journal of Clinical Oncology, 22*(22), 4632-4642.

Dionigi, G., Bianchi, V., Villa, F, Rovera, F, Boni, L, Annoni, M, . . . Dionigi, R. (2007). Differences between familial and sporadic forms of colorectal cancer with DNA microsatellite instability. *Surgical Oncology*, 16, 37-42.

Ferrari, K., Scelfo, A., Jammula, S., Cuomo, A., Barozzi, I., Stützer, A., . . . Pasini, D. (2014). Polycomb-Dependent H3K27me1 and H3K27me2 Regulate Active Transcription and Enhancer Fidelity. *Molecular Cell, 53*(1), 49-62.

Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., … Paul, C. L. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(5), 1827–1831.

Fraga, M., & Esteller, M. (2002). DNA methylation: a profile of methods and applications. *BioTechniques*, 33, 632, 634, 636-49.

Hanahan, D., & Weinberg, R. (2000). The Hallmarks of Cancer. *Cell*, 100, 57-70.

Hanahan , D., & Weinberg , R. (2011). Hallmarks of Cancer: The Next Generation. *Cell*, 144(5), 646-674.

Heng, H., Bremer, S., Stevens, J., Horne, S., Liu, G., Abdallah, B., . . . Ye, C. (2013). Chromosomal instability (CIN): what it is and why it is crucial to cancer. *Cancer metastasis reviews*, 32, 325-40.

Holschneider, C., & Berek, J. (2000). Ovarian cancer: epidemiology, biology, and prognostic factors. *Seminars in surgical oncology*, 19, 3-10.

Hsu, C., Peng, K., Kang, M., Chen, Y., Yang, Y., Tsai, C., . . . Juan, L. (2012). TET1 Suppresses Cancer Invasion by Activating the Tissue Inhibitors of Metalloproteinases. *Cell Reports*, 2, 568-579.

Hughes, L., Khalid-de Bakker, C., Smits, K., van den Brandt, Piet A, Jonkers, Daisy, Ahuja, Nita, . . . van Engeland, Manon. (2013). The CpG island methylator phenotype in colorectal cancer: Progress and problems. Biochimica et Biophysica Acta (BBA) - *Reviews on Cancer*, 1825, 77-85.

Huszar, M., Pfeifer, M., Schirmer, U., Kiefel, H., Konecny, G., Ben-Arie, A., . . . Fogel, Mina. (2009). Up-regulation of L1CAM is linked to loss of hormone receptors and E-cadherin in aggressive subtypes of endometrial carcinomas. *The Journal of Pathology*, 220, 551-561.

Ichimura, N., Shinjo, K., Ohka, F., Katsushima, K., Hatanaka, A., Tojo, M., . . . Kondo, Y. (2015). Abstract 1054: Aberrant TET1 methylation closely associated with CpG island methylator phenotype in colorectal cancer. *Cancer Research Cancer Res, 75*(15 Supplement), 1054-1054.

Isola, J., DeVries, S., Chu, L., Ghazvini, S., & Waldman, F. (1994). Analysis of changes in DNA sequence copy number by comparative genomic hybridization in archival paraffin-embedded tumor samples. *The American Journal of Pathology*, 145(6), 1301–1308.

Jayson, G. C., Kohn, E. C., Kitchener, H. C., & Ledermann, J. A. (2014). Ovarian cancer. *Lancet*, 384(9951), 1376–1388.

Jones, S., Wang, T.-L., Shih, I.-M., Mao, T.-L., Nakayama, K., Roden, R., … Papadopoulos, N. (2010). Frequent Mutations of Chromatin Remodeling Gene ARID1A in Ovarian Clear Cell Carcinoma. *Science (New York, N.Y.)*, 330(6001), 228–231.

Jones, P. A., & Baylin, S. B. (2007). The Epigenomics of Cancer. *Cell*, 128(4), 683-692.

Kato, N., Tamura, G., & Motoyama, T. (2007). Hypomethylation of hepatocyte nuclear factor-1beta (HNF-1beta) CpG island in clear cell carcinoma of the ovary. *Virchows Archiv Virchows Arch, 452*(2), 175-180.

Kato, K., Maesawa, C., Itabashi, T., Fujisawa, K., Otsuka, K., Kanno, S. ... Masuda, T. (2009). DNA hypomethylation at the CpG island is involved in aberrant expression of the L1 cell adhesion molecule gene in colorectal cancer. *International Journal of Oncology*, 35, 467-476.

Kenwrick, S., Watkins, A., & Angelis, E. D. (2000)). Neural cell recognition molecule L1: relating biological complexity to human disease mutations. *Human Molecular Genetics*, 9(6), 879-886.

Ketabi, Z., Bartuma, K., Bernstein, I., Malander, S., Grönberg, H., Björck, E., . . . Nilbert, M. (2011). Ovarian cancer linked to lynch syndrome typically presents as early-onset, non-serous epithelial tumors. *Gynecologic Oncology, 121*(3), 462-465.

Kim, K., Friso, S., & Choi, S.-W. (2009). DNA methylation, an epigenetic mechanism connecting folate to healthy embryonic development and aging. *The Journal of Nutritional Biochemistry*, 20(12), 917–926.

Kurzawski, G., Suchy, J., Debniak, T., Kladny, J., & Lubinski, J. (2004). Importance of microsatellite instability (MSI) in colorectal cancer: MSI as a diagnostic tool. *Annals of Oncology*, 15 Suppl 4, iv283-284.

Lahiri, D. K., & Nurnberger, J. I. (1991). A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. *Nucleic Acids Research*, 19(19), 5444.

Lengauer, C., Kinzler, K., & Vogelstein, B. (1998). Genetic instabilities in human cancers. *Nature*, 396, 643-649.

Li, L., Li, C., Mao, H., Du, Z., Chan, W., Murray, P., . . Ambinder, R. (2016). Epigenetic inactivation of the CpG demethylase TET1 as a DNA methylation feedback loop in human cancers. *Scientific Reports*, 6, 26591.

Li, Y., & Tollefsbol, T. O. (2011). DNA methylation detection: Bisulfite genomic sequencing analysis. *Methods in Molecular Biology (Clifton, N.J.), 791*, 11–21.

Lynch, H., Lynch, P., Lanspa, S., Snyder, C., Lynch, J., & Boland, C. (2009). Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. *Clinical Genetics*, 76, 1–18.

Mao, T.-L., & Shih, I.-M. (2013). The roles of ARID1A in gynecologic cancer. *Journal of Gynecologic Oncology*, 24(4), 376–381.

Marabita, F., Almgren, M., Lindholm, M. E., Ruhrmann, S., Fagerstrom-Billai, F., Jagodic, M., . . . Gomez-Cabrero, D. (2013). An evaluation of analysis pipelines for DNA methylation profiling using the Illumina HumanMethylation450 BeadChip platform. *Epigenetics*, 8(3), 333-346.

Nakamura, K., Banno, K., Yanokura, M., Iida, M., Adachi, M., Masuda, K., . . . Aoki, D. (2014). Features of ovarian cancer in Lynch syndrome (Review). *Molecular and Clinical Oncology*, 2(6), 909-916.

Negrini, S., Gorgoulis, V., & Halazonetis, T. (2010). Genomic instability--an evolving hallmark of cancer. *Nature reviews*. Molecular cell biology, 11, 220-8.

Nemejcova, K., Ticha, I., Kleiblova, P., Bartu, M., Cibula, D., Jirsova, K., & Dundr, P. (2016). Expression, Epigenetic and Genetic Changes of HNF1B in Endometrial Lesions. *Pathology & Oncology Research Pathol. Oncol. Res., 22*(3), 523-530.

Nieminen, T., Niskakoski, A., & Peltomäki, P. (2016). Epigenetic mechanisms in Lynch Syndrome. *eLS. John Wiley & Sons, Ltd, Chichester.*

Nieminen, T. T., Gylling, A., Abdel-Rahman, W. M., Nuorva, K., Aarnio, M., Renkonen-Sinisalo, L., . . . Peltomaki, P. (2009). Molecular Analysis of Endometrial Tumorigenesis: Importance of Complex Hyperplasia Regardless of Atypia. *Clinical Cancer Research, 15*(18), 5772–5783.

Nygren, A. O. H., Ameziane, N., Duarte, H. M. B., Vijzelaar, R. N. C. P., Waisfisz, Q., Hess, C. J., Schouten, JP., Errami, A. (2005). Methylation-Specific MLPA (MS-MLPA):

simultaneous detection of CpG methylation and copy number changes of up to 40 sequences. *Nucleic Acids Research*, 33, e128.

Niskakoski, A., Kaur, S., Staff, S., Renkonen-Sinisalo, L., Lassus, H., Järvinen, H. J., … Peltomäki, P. (2014). Epigenetic analysis of sporadic and Lynch-associated ovarian cancers reveals histology-specific patterns of DNA methylation. *Epigenetics*, *9*(12), 1577–1587.

Peltomäki, P. (2014). Epigenetic mechanisms in the pathogenesis of Lynch syndrome. *Clinical Genetics, 85*(5), 403-412.

Prat, J. (2012). Ovarian carcinomas: five distinct diseases with different origins, genetic alterations, and clinicopathological features. *Virchows Arch*, 460(3), 237-249.

Reisman, D., Glaros, S., & Thompson, E. (2009). The SWI/SNF complex and cancer. *Oncogene*, 28, 1653–1668 .

Riggs, A. D., & Porter, T. N. (1996). Overview of epigenetic mechanisms. *Cold Spring Harbor Monograph Archive*, 32, 29-45.

Ryan, A. J., Susil, B., Jobling, T. W., & Oehler, M. K. (2005). Endometrial cancer. *Cell Tissue Res Cell and Tissue Research, 322*(1), 53-61.

Sameer, A., Nissar, S., & Fatima, K. (2014). Mismatch repair pathway: molecules, functions, and role in colorectal. *European journal of cancer prevention*, 23, 246-57.

Schirmer, U., Fiegl, H., Pfeifer, M., Zeimet, A., Müller-Holzner, E., Bode, P., . . . Altevogt, P. (2013). Epigenetic regulation of L1CAM in endometrial carcinoma: comparison to cancer–testis (CT-X) antigens. *BMC Cancer*, 13, 1-12.

Shames, D., Minna, J., & Gazdar, A. (2007). Methods for detecting DNA methylation in tumors: From bench to bedside. *Cancer Letters*, 251, 187-198.

Sharma, S., Kelly, T. K., & Jones, P. A. (2010). Epigenetics in cancer. *Carcinogenesis*, 31(1), 27-36.

Shen, H., Fridley, B. L., Song, H., Lawrenson, K., Cunningham, J. M., Ramus, S. J., … Pearce, C. L. (2013). Epigenetic analysis leads to identification of *HNF1B* as a subtype-specific susceptibility gene for ovarian cancer. *Nature Communications, 4*, 10.1038/ncomms2629.

Shen, L., & Waterland, R. A. (2007). Methods of DNA methylation analysis. *Current Opinion in Clinical Nutrition and Metabolic Care, 10*(5), 576-581.

Stoffel, E., Mukherjee, B., Raymond, V. M., Tayob, N., Kastrinos, F., Sparr, J., . . . Gruber, S. B. (2009). Calculation of Risk of Colorectal and Endometrial Cancer Among Patients with Lynch Syndrome. *Gastroenterology*, 137(5), 1621-1627.

Terasawa, K., Toyota, M., Sagae, S., Ogi, K., Suzuki, H., Sonoda, T. … Tokino, T. (2006). Epigenetic inactivation of TCF2 in ovarian cancer and various cancer cell lines. *British Journal of Cancer*, 94(6), 914–921.

Tsuchiya, A., Sakamoto, M., Yasuda, J., Chuma, M., Ohta, T., Ohki, M., … Hirohashi, S. (2003). Expression Profiling in Ovarian Clear Cell Carcinoma : Identification of Hepatocyte Nuclear Factor-1β as a Molecular Marker and a Possible Molecular Target for Therapy of Ovarian Clear Cell Carcinoma. *The American Journal of Pathology*, 163(6), 2503–2512.

Umar, A., Risinger, J. I., Hawk, E. T., & Barrett, J. C. (2004). Testing guidelines for hereditary non-polyposis colorectal cancer. *Nature review. Cancer*, 4(2), 153-158.

Vogelstein, B.;& Kinzler, K. (2004). Cancer genes and the pathways they control. *Nature Medicine*, 10, 789-799.

Wang, S., Wang, Z., & Mittal, K. (2015). Concurrent endometrial intraepithelial carcinoma (EIC) and endometrial hyperplasia. *Human Pathology: Case Reports*, 2(1), 1-4.

Wilson, B., & Roberts, C. (2011). SWI/SNF nucleosome remodellers and cancer. *Nature reviews. Cancer*, 11, 481-92.

Wong, J. J., Hawkins, N. J., & Ward, R. L. (2007). Colorectal cancer: a model for epigenetic tumorigenesis. *Gut*, 56(1), 140-148.

Yu, D.-D., Guo, S.-W., Jing, Y.-Y., Dong, Y.-L., & Wei, L.-X. (2015). A review on hepatocyte nuclear factor-1beta and tumor. *Cell & Bioscience*, 5, 58.

Zhang, X., Sun, Q., Shan, M., Niu, M., Liu, T., Xia, B., … Pang, D. (2013). Promoter Hypermethylation of ARID1A Gene Is Responsible for Its Low mRNA Expression in Many Invasive Breast Cancers. *PLoS ONE*, 8(1), e53931.